

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 320.034 Seconds  
(without alignments)  
650.171 Million cell updates/sec

Title: US-10-671-242-43

Perfect score: 2850

Sequence: 1 MAGPPALPPPTAAATAA.....HIKVLQGHFDDPDGFLG 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2850	100.0	538	4	AAB82806	Human low
2	2594	91.0	550	4	AAB82807	Rabbit lo
3	1445.5	50.7	317	4	AAB82798	Rabbit lo
4	1439.5	50.5	317	2	AAW49038	Rabbit lo
5	1260	44.2	241	3	AAB58957	Breast an
6	1191.5	41.8	252	4	AAB82800	Rabbit lo
7	1170	41.1	217	2	AAW49041	Human low
8	1170	41.1	217	4	AAB82803	Human low
9	1086.5	38.1	232	4	AAB82799	Rabbit lo
10	708	24.8	132	7	ADBS9158	Human Pro
11	478	16.8	316	4	AAU31793	Novel hum
12	340	11.9	801	8	ADQ97686	Mouse can
13	336	11.8	783	2	AAW37151	Mouse neu
14	336	11.8	787	2	AAW37152	Mouse neu
15	336	11.8	802	2	AAW37153	Mouse neu
16	336	11.8	802	4	AAU09139	Mammalian
17	324	11.4	397	8	ADQ97684	Mouse can
18	323	11.3	639	8	ADN21309	Bacterial
19	319.5	11.2	783	3	ABO69289	Pseudomon
20	312	10.9	61	3	AAB34365	Human sec
21	311.5	10.9	1151	4	ABBS61598	Drosophil
22	311	10.9	416	5	ABG70108	Human pre
23	311	10.9	505	2	Aaw46889	Human Neu
24	311	10.9	505	2	AAW53808	N-Wiskott
25	311	10.9	505	4	AAG67338	Amino aci

ALIGNMENTS

RESULT 1

AAB82806

ID AAB82806 standard; protein; 538 AA.

XX AC AAB82806;

DT 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 2 (LBP-2).

KW Low density lipoprotein binding protein 2; LBP-2; LDL; human; atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

OS Homo sapiens.

XX WO200164874-A2.

XX PD 07-SEP-2001.

XX PF 28-FEB-2001; 2001WO-US006356.

XX PR 02-MAR-2000; 2000US-00517849.

XX PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX DR N-PSDB; AAH26499.

PT New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence was deduced from the coding region of isolated genomic DNA (see AAH26499). It differs from the sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the presence of an additional 321 amino acids at the N-terminus (the cDNA clone is 5' truncated). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529 -538 (see AAB82803-12) of the present sequence are claimed. Methods of

AAG67360 Amino aci  
AAM52320 Bovine N-W  
AAM52316 Human N-W  
AAE38438 Human N-W  
ABG16505 Novel hum  
AAO82331 MDR rela  
ABU08487 S. pneumo  
AAW46890 Rat Neura  
AAG67355 Amino aci  
AAM52319 Rat N-WAS  
AAE38439 Rat N-WAS  
ABG11919 Novel hum  
AAW72204 HSV-2 str  
AAR60620 Protein f  
ABG03533 Novel hum  
ADCS9465 Novel hum  
ABG09876 Novel hum  
ADC31236 Human nov  
AAG67370 Amino aci

PD - 7/14/2002

Len 4 + W/O

CC determining if an animal is at risk for atherosclerosis, methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed  
 XX  
 SQ Sequence 538 AA;

Query Match 100.0%; Score 2850; DB 4; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-165;  
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPPALPPPTAAATTAAASSSSAASPHYQEWILDTIDSLRSKARPDLEICRMVR 60  
 DB 1 MAGPPALPPPTAAATTAAASSSSAASPHYQEWILDTIDSLRSKARPDLEICRMVR 60  
 QY 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
 DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
 QY 121 AAAAAAPPTTAPPPPPAPVAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180  
 DB 121 AAAAAAPPTTAPPPPPAPVAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180  
 QY 181 PPAPAPPAVAPPAGPRRAPPAAVAAAREPLPPPPPPAPPOQOQPPPPPPPPPPGGAV 240  
 DB 181 PPAPAPPAVAPPAGPRRAPPAAVAAAREPLPPPPPPAPPOQOQPPPPPPPPPPGGAV 240  
 QY 241 RAGGAARPVSLREVVRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGALALPRG 300  
 DB 241 RAGGAARPVSLREVVRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLERTRLGALALPRG 300  
 QY 301 DRPGRAPPAASARPSKRGGEERVLKEEEDDDDEDEEDDVSGSEVPESDRPAGAQ 360  
 DB 301 DRPGRAPPAASARPSKRGGEERVLKEEEDDDDEDEEDDVSGSEVPESDRPAGAQ 360  
 QY 361 HHQNGERGQSAKERVKWTPCGPHQGDGGRGAPGSGTRQVFSMAANKEGGTASVA 420  
 DB 361 HHQNGERGQSAKERVKWTPCGPHQGDGGRGAPGSGTRQVFSMAANKEGGTASVA 420  
 QY 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKGKEKSDPVWVTMDVVEYFTGAPPEQA 480  
 DB 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKGKEKSDPVWVTMDVVEYFTGAPPEQA 480  
 QY 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538  
 DB 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538

## RESULT 2

AAB82807  
 ID AAB82807 standard; protein; 550 AA.

XX AAB82807;  
 AC  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Oryctolagus cuniculus.  
 OS  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX

PA (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;  
 XX WPI; 2001-565505/63.  
 DR N-PSDB; AAH26500.  
 XX

PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 PS Claim 13(k); Fig 2a; 143pp; English.  
 XX

CC The present sequence is that of a full-length sequence of novel rabbit  
 CC low density lipoprotein binding protein 2 (LBP-2). The amino acid  
 CC sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit  
 CC LBP-2 is an example of claimed polypeptides of the invention, termed  
 CC LBP-2, that are capable of binding to native and methylated low density  
 CC lipoproteins. Also claimed are biologically active fragments and  
 CC analogues of LBPs, polynucleotides encoding LBPs, as well as expression  
 CC vectors, cells and methods of producing the LBPs. Polypeptides having  
 CC amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-  
 CC 18) of the present sequence are claimed. Methods for determining if an  
 CC animal is at risk for atherosclerosis, methods for evaluating an agent  
 CC for use in treating atherosclerosis, and methods for treating a cell  
 CC having an abnormality in structure or metabolism of LBP are also claimed,  
 CC as are pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions  
 XX

SQ Sequence 550 AA;

Query Match 91.0%; Score 2594; DB 4; Length 550;

Best Local Similarity 91.1%; Pred. No. 1.3e-149;  
 Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPTAAATT-AAAASSAASPHYQEWILDTIDSLRSKARPDLEICRMVR 59  
 DB 1 MAGPPALPPPTAAATTAAAAASSSSAASPHYQEWILDTIDSLRSKARPDLEICRMVR 60  
 QY 60 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 119  
 DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
 QY 120 AAAAAAPPTTAPPPPPAPV-AAAAAPARAPR-AAAAAATAPSPGPAQPGPRAQRAAP 176  
 DB 121 --AAAAAPPTTAPPPPPAPVAAAPARAPRAAAAAAATAPSPGPAQPGPRAQRAAP 178  
 QY 177 LAAPPAPAPPAVAPPAGPRRAPP--AAVAREPPLPPPPPPPPPPPPPPPPPPPPPP 229  
 DB 179 LAAPPAPAPPAVAPPAGPRRAPPAAVAREPPLPPPPPPPPPPPPPPPPPPPPPP 238  
 QY 230 QPQPPPEGGAAGGAAARPVSLREVVRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLER 289  
 DB 239 QPQPPPEGGAAGGAAARPVSLREVVRVYLGSGGAGGLTRGRVQGLLEEEAAARGRLER 298  
 QY 290 TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLKEEEDDDDEDEDEDEDD--VSEG 347  
 DB 299 TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLKEEEDDDDEDEDEDEDD 358  
 QY 348 SEVPESDRPAGAQHHQLN-GERGQSAKERVKWTCPGPHQGDGGRGAPGSGTRQVFS 406  
 DB 359 SEVPESDRPAGAQHHQLNGERGTPQAKERAKESWLCGPHQGEGRGPAAGSGTRQVFS 418  
 QY 407 MAANKEGGTASVATGPDSPSPVPLPGKALPGADGTPGCPGPKGKEKSDPVWVTMD 466  
 DB 419 MAALSKEGGSASTTGPDSPSPVPLPGKALPGADGTPGCPGPKGKEKSDPVWVTMD 478  
 QY 467 VVEYFTGAPPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQ 526  
 DB 479 VVEYFTGAPPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQ 538  
 QY 527 HFEDDDPDGFLG 538  
 DB 539 HFEDDDPDGFLG 550

RESULT 3  
ID AAB82798  
XX AAB82798 standard; protein; 317 AA.  
AC AAB82798;  
XX  
DT 12-NOV-2001 (first entry)  
XX  
DE Rabbit low density lipoprotein binding protein 2 (LBP-2).  
XX  
KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
XX  
OS Oryctolagus cuniculus.  
XX  
PH Key Location/Qualifiers  
FT Misc-difference 10 /note= "encoded by TAG"  
FT  
PN WO200164874-A2.  
XX  
XX 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US006356.  
XX  
PR 02-MAR-2000; 2000US-00517849.  
PR 14-JUL-2000; 2000US-00616289.  
XX  
PA (BOST-) BOSTON HEART FOUND INC.  
XX  
PI Lees AM, Lees RS, Law SW, Arjona AA;  
XX  
DR WPI; 2001-565505/63.  
DR N-PSDB; AAB26488.  
XX  
XX New isolated low density lipoprotein binding polypeptide for treating,  
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
XX  
XX Claim 13(b); Fig 2b; 143pp; English.  
XX  
CC The present sequence is that of a partial sequence of novel rabbit low  
CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
CC deduced from an isolated cDNA clone (see AAB26488). Full-length rabbit  
CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed  
CC polypeptides of the invention, termed LBPs, that are capable of binding  
CC to native and methylated low density lipoproteins. Also claimed are  
CC biologically active fragments and analogues of LBPs, polynucleotides  
CC encoding LBPs, as well as expression vectors, cells and methods of  
CC producing the LBPs. Methods of determining if an animal is at risk for  
CC atherosclerosis, methods for evaluating an agent for use in treating  
CC atherosclerosis, and methods for treating a cell having an abnormality in  
CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
XX compositions  
SQ Sequence 317 AA;  
  
Query Match 50.7%; Score 1445.5; DB 4; Length 317;  
Best Local Similarity 89.3%; Pred. No. 4.8e-80;  
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;  
  
QY 235 PEGGAVRAGGAAPVSLREVYVYLGSGGAGGLTGRVQGLLEEEAARGRLERTLGA 294  
Db 11 PKGAARAGGAPVSLREVYVYLGSGGAGGLTGRVQGLLEEEAARGRLERTLGA 70  
  
QY 295 LALPRGDRPGRAPPAASAPSRKSGGEERVLKEEEDDEDEDEDD--VSEGSEVPE 352  
Db 71 LALPRGDRPGRAPPAASAPSRKSGGEERVLKEEEDDEDEDEDD--VSEGSEVPE 130  
  
QY 353 SDRPAGAQQHQLN-GERGPOSARKERVKWTFCPHQGDGEGRAPGSGTRQVFSMAAN 411  
|||||

Db 131 SDRPAGAQQHQLNGSGRGPGQTAKERAKWSLCGPHFGQBEGRPAAGSGTRQVFSMAALS 190  
QY 412 KEGGTASVATGPDSPSPVPLPGCKPALPGADGTFPGCPGRKEKPSDPVETWMDVVEYF 471  
Db 191 KEGGSASSTGPDSPSPVPLPGCKPALPGADGTFPGCPGRKEKPSDPVETWMDVVEYF 250  
QY 472 TEAGPPEQATAFQEQEIDGKSLLLMORTDVLTLGLSIRLPALKIYEHKIKVLOQGHFEDD 531  
Db 251 TEAGPPEQATAFQEQEIDGKSLLLMORTDVLTLGLSIRLPALKIYEHKIKVLOQGHFEDD 310  
QY 532 DPDGFLG 538  
Db 311 DPEGFLG 317  
  
RESULT 4  
AAB49038  
ID AAB49038 standard; protein; 317 AA.  
XX  
AC AAB49038;  
XX  
DT 09-NOV-1998 (first entry)  
XX  
DE Rabbit low density lipoprotein binding protein LBP-2.  
XX  
KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.  
XX  
OS Oryctolagus cuniculus.  
XX  
PH Key Location/Qualifiers  
FT Misc-difference 10 /note= "encoded by TAG"  
FT Protein 66..317 /note= "Claim 1"  
FT Protein 86..317 /note= "Claim 1"  
FT Peptide 105..132 /note= "Claim 2"  
FT Peptide 105..120 /note= "Claim 2"  
FT Peptide 121..132 /note= "Claim 2"  
FT Peptide 211..220 /note= "Claim 2"  
XX  
PN WO9823282-A1.  
XX  
PD 04-JUN-1998.  
XX  
PF 26-NOV-1997; 97WO-US021857.  
XX  
PR 27-NOV-1996; 96US-0031930P.  
PR 03-JUN-1997; 97US-0048547P.  
XX  
PA (BOST-) BOSTON HEART FOUND INC.  
XX  
PI Lees AM, Lees RS, Law SW, Arjona AA;  
XX  
DR WPI; 1998-322455/28.  
DR N-PSDB; AAV32835.  
XX  
PT Nucleic acid encoding low density lipoprotein binding proteins and  
PT related vectors - transformed cells, proteins, and modulators of binding,  
PT useful for treatment and diagnosis of atherosclerosis and for identifying  
PT subjects at risk.  
XX  
PS Claim 1; Fig 2; 47pp; English.  
XX  
CC This polypeptide comprises novel rabbit low density lipoprotein (LDL)  
CC binding protein LBP-2 that is capable of binding both native and methyl  
CC LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA  
CC (see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and





```
XX PN WO200164874-A2.
XX KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
XX PD receptor; human; atherosclerosis; diagnosis; therapy; vaccine.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 8..33
XX FT /note= "Claim 2"
XX FT Peptide 8..22
XX FT /note= "Claim 2"
XX FT Peptide 23..33
XX FT /note= "Claim 2"
XX FT Peptide 208..217
XX FT /note= "Claim 2"
XX PN WO9823282-A1.
XX XX
XX PD 04-JUN-1998.
XX XX
XX PF 26-NOV-1997; 97WO-US021857.
XX XX
XX PR 27-NOV-1996; 96US-0031930P.
XX PR 03-JUN-1997; 97US-0048547P.
XX XX
XX PA (BOST-) BOSTON HEART FOUND INC.
XX XX
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX DR WPI; 1998-322455/28.
XX DR N-PSDB; AAV32838.
XX XX
XX PT Nucleic acid encoding low density lipoprotein binding proteins and
XX PT related vectors - transformed cells, proteins, and modulators of binding,
XX PT useful for treatment and diagnosis of atherosclerosis and for identifying
XX PT subjects at risk.
XX XX
XX PS Claim 1; Fig 7; 47pp; English.
XX XX
XX CC This polypeptide comprises novel human low density lipoprotein (LDL)
XX CC binding protein LBP-2 that is capable of binding both native and methyl
XX CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
XX CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human
XX CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP
XX CC metabolism or structure is diagnostic of a risk for atherosclerosis. The
XX CC invention provides methods for determining if an animal is at risk for
XX CC atherosclerosis (e.g. for prenatal screening); methods for treating
XX CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
XX CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
XX CC methods for treating a cell having an abnormality in LBP structure or
XX CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
XX CC well as recombinant vectors and host cells used to produce recombinant
XX CC LBP
XX XX
XX SQ Sequence 217 AA;
Query Match 41.1%; Score 1170; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e-63;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 322 EERVLEKEEEEDDEDEDDVSEGVSESDRPAQAHHQLNGRGPQSAKERVKEWT 381
Db 1 EERVLEKEEEEDDEDEDDVSEGVSESDRPAQAHHQLNGRGPQSAKERVKEWT 60
QY 382 PCGPHQGDGGRGAPGSGTRQVFSMAANKGGTASVATGDPSPVPLPGKPALPGA 441
Db 61 PCGPHQGDGGRGAPGSGTRQVFSMAANKGGTASVATGDPSPVPLPGKPALPGA 120
QY 442 DGTFFCGPPGRKEKPSDPVETWMDVVEYFTTAGPPEQATAFQEQIDGKSLLMORTDV 501
Db 121 DGTFFCGPPGRKEKPSDPVETWMDVVEYFTTAGPPEQATAFQEQIDGKSLLMORTDV 180
QY 502 LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538
XX XX
XX SQ Sequence 252 AA;
Query Match 41.8%; Score 1191.5; DB 4; Length 252;
Best Local Similarity 88.5%; Pred. No. 1e-64;
Matches 223; Conservative 17; Mismatches 9; Indels 3; Gaps 2;
QY 290 TRIGALALPRGDRPAPPAAASARPSKRGSEVLEKEEEEDDEDEDD--VSEG 347
Db 1 TRIGALALPRGDRPAPPAAASARAARKAGEVLEKEEEEDDEDDDDVVSEG 60
QY 348 SEVPSDRPAGAHHQLN-CERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFS 406
Db 61 SEVPSDRPAGAHHQLNGRGERGPOTAKERAKEWSLGGPHGQEGRGPAGSGTRQVFS 120
QY 407 MAAMNKEGGTASVATGDPSPVPLPGKPALGADGTPGCPGRKEKPSDPVETWMD 466
Db 121 MAALSKEGGASSTTGDPSPVPLPGKPALGADGTPGCPGRKEKPSDPVETWMD 180
QY 467 VVEYFTAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLGPALKIYEHHIKVLQOG 526
Db 181 VVEYFTAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLGPALKIYEHHIKVLQOG 240
QY 527 HFEDDDPDGFLG 538
Db 241 HFEDDDPDGFLG 252
XX XX
XX RESULT 7
XX ID AAW49041
XX ID AAW49041 standard; protein; 217 AA.
XX AC AAW49041;
XX XX
XX DT 09-NOV-1998 (first entry)
XX XX
XX DE Human low density lipoprotein binding protein LBP-2.
```

Db 181 LTGLSIRLGPALKIYEHKIKVLQGGHFEDDDPDGLG 217

RESULT 8  
 AAB82803  
 ID AAB82803 standard; protein; 217 AA.  
 AC AAB82803;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Human low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 PA (BOST-) BOSTON HEART FOUND INC.  
 XX  
 PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX  
 DR WPI; 2001-565505/63.  
 DR N-PSDB; AAH26494.  
 XX  
 PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 PS Claim 13(g); Fig 7B; 143pp; English.

XX The present sequence is that of the N-terminal portion of novel human low  
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
 CC deduced from an isolated partial cDNA clone (see AAH26494). A full-length  
 CC sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP  
 CC polypeptides of the invention that are capable of binding to native and  
 CC methylated low density lipoproteins. Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding the LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs. Methods  
 CC of determining if an animal is at risk for atherosclerosis, methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed  
 XX  
 SQ Sequence 217 AA;

Query Match 41.1%; Score 1170; DB 4; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-63;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 EERVLSKEEEDDEDEDDYSEGSEVPESDRPAGAQHQLNGRGPQSAKERVKEWT 381  
 Db 1 EERVLSKEEEDDEDEDDYSEGSEVPESDRPAGAQHQLNGRGPQSAKERVKEWT 60  
 QY 382 PCGPHQOQDEGRGPAFCSGTRQVFSMAAMNKEGGTASVATGPDSPFVPLPGKPALPGA 441  
 Db 61 PCGPHQOQDEGRGPAFCSGTRQVFSMAAMNKEGGTASVATGPDSPFVPLPGKPALPGA 120  
 QY 442 DGTFFCGPKRKEKPSDPVWTVMDVVEYFTAGFPEQATAFQEQIDGKSLLLMORTDV 501  
 Db 121 DGTFFCGPKRKEKPSDPVWTVMDVVEYFTAGFPEQATAFQEQIDGKSLLLMORTDV 180  
 QY 502 LTGLSIRLGPALKIYEHKIKVLQGGHFEDDDPDGLG 538

Db 181 LTGLSIRLGPALKIYEHKIKVLQGGHFEDDDPDGLG 217

RESULT 9  
 AAB82799  
 ID AAB82799 standard; protein; 232 AA.  
 AC AAB82799;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Oryctolagus cuniculus.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 PA (BOST-) BOSTON HEART FOUND INC.  
 XX  
 PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX  
 DR WPI; 2001-565505/63.  
 DR N-PSDB; AAH26489.  
 XX  
 PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 PS Claim 13(c); Fig 3; 143pp; English.

XX The present sequence is that of a partial sequence of novel rabbit low  
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
 CC deduced from an isolated cDNA clone (see AAH26489). Full-length rabbit  
 CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed  
 CC polypeptides of the invention, termed LBPs, that are capable of binding  
 CC to native and methylated low density lipoproteins. Also claimed are  
 CC biologically active fragments and analogues of LBPs, polynucleotides  
 CC encoding LBPs, as well as expression vectors, cells and methods of  
 CC producing the LBPs. Methods of determining if an animal is at risk for  
 CC atherosclerosis, methods for evaluating an agent for use in treating  
 CC atherosclerosis, and methods for treating a cell having an abnormality in  
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
 CC compositions  
 XX  
 SQ Sequence 232 AA;

Query Match 38.1%; Score 1086.5; DB 4; Length 232;  
 Best Local Similarity 87.5%; Pred. No. 2.3e-58;  
 Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;

QY 310 ASARPSRKRGGEFVLEKEEEDDEDEDD--VSEGSEVPESDRPAGAQHQLN-G 366  
 Db 1 ASARARNKRAEERVLEKEEEDDEDEDDVVSEGSEVPESDRPAGAQHQLNGG 60  
 QY 367 ERGPQSAKERVKEWTPCGPHQOQDEGRGPAFCSGTRQVFSMAAMNKEGGTASVATGPDSP 426  
 Db 61 ERGPQTAKEAKESLTCGPHQOQDEGRGPAFCSGTRQVFSMAALSKEGGSASTTGPDS 120  
 QY 427 SPVPLPPGKPALPGADGTPFGCPGPKRKEKPSDPVWTVMDVVEYFTAGFPEQATAFQEQ 486  
 Db 121 SPVPLPPGKPALPGADGTPFGCPGPKRKEKPSDPVWTVMDVVEYFTAGFPEQATAFQEQ 180

Qy 487 BIDGKSLLMQRTDVLGSLRGLPALKIYEHKIKVLQOQHFDDEDDPDGFLG 538  
Db 181 BIDGKSLLMQRTDVLGSLRGLPALKIYEHKIKVLQOQHFDDEDDPDGFLG 232

RESULT 10  
ADE59158  
ID ADE59158 standard; protein; 132 AA.

AC ADE59158;  
XX  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Human Protein XP\_031299, SEQ ID NO 5049.  
DE Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
XX Homo sapiens.

OS  
XX  
XX WO2003016475-A2.  
XX  
XX 27-FEB-2003.  
XX

XX 14-AUG-2002; 2002WO-US025765.  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346392P.  
PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;  
XX

PI WPI; 2003-269312/26.  
XX GENBANK; XP\_031299.

DR New composition comprising two or more isolated polypeptides, useful for  
DR preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

PS The invention discloses a composition comprising two or more isolated rat  
XX or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 132 AA;

Query Match 24.8%; Score 708; DB 7; Length 132;  
Best Local Similarity 100.0%; Pred. No. 1.3e-35;  
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 MAAMNKEGGTASVATGPDSPVPLPGKPALPGADGTPFGCPGRKEKPSDPVETVMD 466  
Db 1 MAAMNKEGGTASVATGPDSPVPLPGKPALPGADGTPFGCPGRKEKPSDPVETVMD 60  
Qy 467 VVEYFTGAGPPEQATAFQEQEIDGKSLLMQRTDVLGSLRGLPALKIYEHKIKVLQOQ 526  
Db 61 VVEYFTGAGPPEQATAFQEQEIDGKSLLMQRTDVLGSLRGLPALKIYEHKIKVLQOQ 120  
Qy 527 HFEDDDPDGFLG 538  
Db 121 HFEDDDPDGFLG 132

RESULT 11

AAU31793  
ID AAU31793 standard; protein; 316 AA.

XX  
XX  
XX AAU31793;  
XX

XX 18-DEC-2001 (first entry)  
XX

XX Novel human secreted protein #2284.  
XX

XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

XX 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX vaccination, testing and therapy.

XX Claim 20; Page 509; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent or  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
XX AAU29510-AAU3304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention

XX Sequence 316 AA;

Query Match 16.8%; Score 478; DB 4; Length 316;  
Best Local Similarity 38.5%; Pred. No. 3e-21;  
Matches 136; Conservative 9; Mismatches 84; Indels 124; Gaps 14;

QY 235 PEGAVRAGGAARPVSLREYVYLGSGGAGGLTRGRVQGLLEEEAARGRLERLGA 294  
DB 2 PRAGNAGGG-----GGRPSGAHPSR-----SA 25  
QY 295 LALPRGRDPRGAPPAASARSRKRGGEERVLKKEEEDD-DEDEEDDVSGSV-PE 352  
DB 26 CGCPAGTRPGRAPPAASARSRKRGGEERVLKKEEEDDDEKXRRXRRSMCQRSEVAPS 85  
QY 353 SDRPAGAHQHLNGERGQSAKVRKWTGCG---PHQODEG-RGPAFGSGTRQVFSM 407  
DB 86 SDRPGRCP-----APPSL-----TASGDLRVRRGRSGPPGTAPGPGXAGASP 130  
QY 408 AANNKEGGT-----ASVATGDSPPVPLPPGKPALPGADGTPFGCPP 450  
DB 131 GQRHPPGCSWPQYXTRKGEQVFFSPASVATGDSPPVPLPPGKPALPGADGTPFGCPP 190  
QY 451 GRKEKSDP-VETVMDV-VETTERAGFPQATAPQEQEIDGKSLLLMORTDVLTLGLSTR 508  
DB 191 GRKEKFPDRSSWNVMGFRSNILLEAXFPQATAP-----R 227  
QY 509 LGPALKIYEH-----IKVLQGHFEDD-----DPDGFL 537  
DB 228 AGNWXQIFAHAHRCASPCSPRASPENIRAPHQAGASARPLXGMWDPDGFL 280

RESULT 12  
ADQ97686  
ID ADQ97686 standard; protein; 801 AA.  
XX ADQ97686;  
XX 07-OCT-2004 (first entry)  
XX Mouse cancer associated sequence MP2-10-028, SEQ ID 663.  
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.  
XX Mus musculus.  
XX WO2004060304-A2.  
XX 22-JUL-2004.  
XX 22-DEC-2003; 2003WO-US041389.  
XX 27-DEC-2002; 2002US-00330773.  
XX (SAGR-) SAGRES DISCOVERY INC.  
XX Morris DW, Malandro MS;  
XX WPI; 2004-543781/52.  
XX New isolated cancer associated nucleic acids comprising at least 10  
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating  
XX cancers such as leukemia and lymphoma.  
XX Claim 1; SEQ ID NO 663; 199pp; English.  
XX The present invention relates to cancer associated sequences (ADQ97025-  
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
XX data for this patent did not form part of the printed specification, but  
XX was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 801 AA;  
SQ

Query Match 11.9%; Score 340; DB 8; Length 801;  
Best Local Similarity 26.6%; Pred. No. 1.9e-12;  
Matches 123; Conservative 40; Mismatches 175; Indels 124; Gaps 18;

QY 4 PPALPP-PETAAAAATTAAASSSA---ASPHQEWILDTIDSLSRKARP-----50  
DB 308 PSTPTPTPLRHAATRFATSLGSAFHPVLPHYA---TVPRLNKNSRSPSVNTPSSQ 362  
QY 51 --LERICRMVRRRHGPEPERTAELE---KLIQRAVLRYKSGISYRNAARVOPPPR 104  
DB 363 PPAKSCAWPTSNFSPPLPSPIMISSPPGKATGPRVLPVCVSPV-----PQM 412  
QY 105 GATPPAPPA-----PRGAPAAAAAAPPPTAPPPPPAPVAAAAAPARAAAAATA 158  
DB 413 PPSPTAPNGSLDSVTVYVSPPTSGPAADPPPPPPPPPPPPPPPPPPPPPLPLPLASLHCGS 472  
QY 159 PPSGPAQP-----GPRQRAAPLAAP-----180  
DB 473 QASPPGTELASTPPSKPSVLPSPSAAAPASAEETPLNPELGDSSASEPGLQAASQPAESP 532  
QY 181 -----PPAPAAP-----AVAPPAGPRAPPAPPAVAREPPLPPPPPPAPPQ 222  
DB 533 TPQGLVLGPPAPPPPPPLPSGPAYASALPPPPG-----PPPPPLPSTGPPPPPPPPPP 590  
QY 223 QQQPPPPQPPPEGGAVRAGAA---RPVS-----LREVRYLGGSGGAGRLT 269  
DB 591 APPPPPPPPAPPLPAGSIFSGSTSEDNRPLTGLAAIAGAKLRKVSIVEDGSPFGG--N 648  
QY 270 RGRVQGLLEEEAARGRLERLIGALALPRGDRPCRPAPPAASARPSRSKRGGEE-RVLEK 328  
DB 649 TGSVS--LASSKADAGR-----GNGLPLGG--SGLMEEMSALLARRRRIAEKGSTIET 698  
QY 329 EEEEDDEDEDEDDVSESGSEVPESDRPAGAHQHLNGERP 370  
DB 699 EQEDRNEDAEPIITAKAPSTSTPEPTKRPWERTNTMNGSKSP 740

RESULT 13  
AAW37151  
ID AAW37151 standard; protein; 783 AA.  
XX AAW37151;  
XX 06-JUL-1998 (first entry)  
XX Mouse neural Mena+ protein.  
XX Neural Mena+ protein; mammalian Ena; Enabled protein; Evi protein;  
XX cytoskeleton; cell morphology; cell adhesion; cell differentiation;  
XX cell growth; cell motility; mouse.  
XX Mus musculus.  
XX Key Location/Qualifiers  
XX Misc-difference 378 /note= "encoded by GGN"  
XX WO9801755-A1.  
XX 15-JAN-1998.  
XX 03-JUL-1997; 97WO-US011669.  
XX 05-JUL-1996; 96US-00675815.  
XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.  
XX (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MEH.  
XX Gertler FB, Soriano P, Wehland J, Niebuhr K;  
XX WPI; 1998-101197/09.  
XX N-PSDB; AAV02998.  
XX

PT Detection of modulators of Mena and Ena-VASP-like genes and proteins -  
 PT used in control of cytoskeletal dynamic events in normal and abnormal  
 XX cell morphology, adhesion, motility, growth and differentiation.  
 PS Example 4; Page 58-60; 77pp; English.  
 XX  
 CC This protein comprises novel murine neural Mena+. Its amino acid sequence  
 CC was deduced from a cDNA clone (see AAV02998) obtained from a mouse brain  
 CC cDNA library. Neural Mena+ contains an exon that introduces 244 amino  
 CC acids between amino acids 238 and 239 of mammalian Ena (Mena, see  
 CC AAW37148). Two other isoforms, neural Mena++ (see AAW37152) and neural  
 CC Mena+++ (see AAW37153), are also disclosed. Unlike Mena, neural Mena  
 CC isoforms exhibit neural tissue-specific distribution. Based on the  
 CC disclosed Mena and Evi genes (see also AAV02996-97) and proteins (see  
 CC also AAW37148-49), a variety of methods and compositions are provided for  
 CC screening, isolating and characterizing endogenous and exogenous factors,  
 CC drugs and therapeutic agents useful to evaluate and/or control  
 CC cytoskeletal dynamic events involved in normal and abnormal cell  
 CC morphology, adhesion, motility, growth and/or differentiation. A method  
 CC of detecting a modulator of Mena activity/expression is claimed  
 XX  
 SQ Sequence 783 AA;

Query Match 11.8%; Score 336; DB 2; Length 783;  
 Best Local Similarity 26.8%; Pred. No. 3.3e-12;  
 Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;  
 QY 4 PPALPP-PETAAATTAATAASSSA---ASPHYQEWILDTIDLSRSRKARPD----- 50  
 DB 290 PTSTPTPLRLHAATRTATSLGSAFHVLPHYA-----TVPRPLNKNRSPSPVNTPSQ 344  
 QY 51 --LERICRMVRRRHGPEPERTRAELE---KLIQRAVLRSYKGSISYRNAARVQPPRR 104  
 DB 345 PPAKSCAMPTSNFSLPPSPIMISSPPGKATGPRPVLPCVSSPV-----PQM 394  
 QY 105 GATPPAPPRA-----PRGAPAAAAAAPPPTPAPPPAPVAAAAAPARAPRAA-----AA 154  
 DB 395 PPSPTAPNGSLDSVTYVPSPPPTSGPAAPPPPPPPPPPPPPPPPPPPPLPLPLASLSHSGS 454  
 QY 155 AATAPPS-----PGP-----AQPGRAPRAAPLAA 179  
 DB 455 QASPPPGTPLASTTSSKPSVLPSPSAGAPASAEPTLNPGLGSSASEPGLQA-ASQPAES 513  
 QY 180 P-----PPAAPAPP-----AVAPPAGRRAPPVAAAREPPLPPPPPPAPP 221  
 DB 514 PTQGLVLPAPPPPPPLPSGPAYASALPPPGP--PPPPPLPSTGTPPPPPPPPLPN 571  
 QY 222 QQQQPPPPQPPPEGGAVRAGGAA---RPVS-----LREVVYLGSGGAGGRL 268  
 DB 572 QAPPPPPPPAPPLPASGIFSGTSDNRLTGLAAAIAGAKLRKVSERVEDGSFPGG-- 629  
 QY 269 TRGRVQGLLEEEAARGRLERLGLALPRGDRPGRAPPAASARPSRKGEE--RVLE 327  
 DB 630 NTGSVS--LASSKADAGR-----GNGPLFGG--SGLMEEMSALLARRRIAEKGSTIE 679  
 QY 328 KEEEDDDDEDEDDVSEGSVEPESDRPAGAQQHQLNGERGP 370  
 DB 680 TEQEDRNEADPEITAKAPSTSTPEPTRKPKWERTNTWNGSKSP 722

RESULT 14  
 AAW37152  
 ID AAW37152 standard; protein; 787 AA.  
 XX  
 AC AAW37152;  
 XX  
 XX 06-JUL-1998 (first entry)  
 DT  
 XX Mouse neural Mena++ protein.  
 DE  
 XX Neural Mena++ protein; mammalian Ena; Enabled protein; Evi protein;  
 KW cytoskeleton; cell morphology; cell adhesion; cell differentiation;  
 KW cell growth; cell motility; mouse.

XX Mus musculus.  
 OS WO9801755-A1.  
 PN 15-JAN-1998.  
 XX  
 PD 03-JUL-1997; 97WO-US0111669.  
 XX  
 PF 05-JUL-1996; 96US-00675815.  
 XX  
 PR (HUTC-) HUTCHINSON CANCER RES CENT FRED.  
 PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.  
 XX  
 PI Gertler FB, Soriano P, Wehland J, Niebuhr K,  
 XX WPI; 1998-101197/09.  
 DR  
 XX Detection of modulators of Mena and Ena-VASP-like genes and proteins -  
 PT used in control of cytoskeletal dynamic events in normal and abnormal  
 PT cell morphology, adhesion, motility, growth and differentiation.  
 PT  
 XX Example 4; Page 60-63; 77pp; English.  
 PS  
 XX This protein comprises novel murine neural Mena+. Its amino acid  
 CC sequence was deduced from a cDNA clone obtained from a mouse brain cDNA  
 CC library. Two other isoforms, neural Mena+ (see AAW37151) and neural  
 CC Mena+++ (see AAW37153), are also disclosed. Unlike mammalian Ena (Mena,  
 CC see AAW37148), neural Mena isoforms exhibit neural tissue-specific  
 CC distribution. Based on the disclosed Mena and Evi genes (see also  
 CC AAV02996-98) and proteins (see also AAW37148-49), a variety of methods  
 CC and compositions are provided for screening, isolating and characterising  
 CC endogenous and exogenous factors, drugs and therapeutic agents useful to  
 CC evaluate and/or control cytoskeletal dynamic events involved in normal  
 CC and abnormal cell morphology, adhesion, motility, growth and/or  
 CC differentiation. A method of detecting a modulator of Mena  
 CC activity/expression is claimed  
 XX  
 SQ Sequence 787 AA;

Query Match 11.8%; Score 336; DB 2; Length 787;  
 Best Local Similarity 26.8%; Pred. No. 3.3e-12;  
 Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;  
 QY 4 PPALPP-PETAAATTAATAASSSA---ASPHYQEWILDTIDLSRSRKARPD----- 50  
 DB 294 PTSTPTPLRLHAATRTATSLGSAFHVLPHYA-----TVPRPLNKNRSPSPVNTPSQ 348  
 QY 51 --LERICRMVRRRHGPEPERTRAELE---KLIQRAVLRSYKGSISYRNAARVQPPRR 104  
 DB 349 PPAKSCAMPTSNFSLPPSPIMISSPPGKATGPRPVLPCVSSPV-----PQM 398  
 QY 105 GATPPAPPRA-----PRGAPAAAAAAPPPTPAPPPAPVAAAAAPARAPRAA-----AA 154  
 DB 399 PPSPTAPNGSLDSVTYVPSPPPTSGPAAPPPPPPPPPPPPPPPPPPPPLPLPLASLSHSGS 458  
 QY 155 AATAPPS-----PGP-----AQPGRAPRAAPLAA 179  
 DB 459 QASPPPGTPLASTTSSKPSVLPSPSAGAPASAEPTLNPGLGSSASEPGLQA-ASQPAES 517  
 QY 180 P-----PPAAPAPP-----AVAPPAGRRAPPVAAAREPPLPPPPPPAPP 221  
 DB 518 PTQGLVLPAPPPPPPLPSGPAYASALPPPGP--PPPPPLPSTGTPPPPPPPPLPN 575  
 QY 222 QQQQPPPPQPPPEGGAVRAGGAA---RPVS-----LREVVYLGSGGAGGRL 268  
 DB 576 QAPPPPPPPAPPLPASGIFSGTSDNRLTGLAAAIAGAKLRKVSERVEDGSFPGG-- 633  
 QY 269 TRGRVQGLLEEEAARGRLERLGLALPRGDRPGRAPPAASARPSRKGEE--RVLE 327  
 DB 634 NTGSVS--LASSKADAGR-----GNGPLFGG--SGLMEEMSALLARRRIAEKGSTIE 683  
 QY 328 KEEEDDDDEDEDDVSEGSVEPESDRPAGAQQHQLNGERGP 370

Db 684 TEQKEDRNEDAEPIITAKAPSTSTPEPTRKWPWERTNTMNGSKSP 726  
|::|||::||: : : || | : : : || :  
155 AATAPPS-----PGP-----AQCGRRAQRAAPLAA 179  
474 QASPPPGTTLASTPSKPSVLPSPSAGAPASAEPTLNPELGSSASEPLQA-ASQPAES 532  
180 P-----PPAPAAAPP-----AVAPPAGRRRAPPPAVAAAREPPLPPPPQPPAPP 221  
533 PTPQGLVLGPPAPPPPPPLPSGPAYASALPPPPGP--PPPPPLPSTGPPPPPPPPPLEN 590  
222 QQQQPPPPQPPPPPEGGAVRAGGAA---RPVS-----LREVVRVLGSGGAGGRL 268  
591 QAPPPPPPPAPPLPASGIFSGSTSDNRPLTGLAAATAGAKLRKVSRYVEDGSFPGGG-- 648  
269 TGRVQGLLEEEAAARGRLERTRIGALALPRGDRPGRAPPAASARPSRSKRGEE-RVLE 327  
649 NTGSVS--LASSKADAGR-----GNGPLPLGG--SGLMEEMSAALLARRRRIAEKGSTIE 698  
328 KEEEDDDDEDEDEDDVSEGSVPESDRPAGAHQLNGERGP 370  
699 TEQKEDRNEDAEPIITAKAPSTSTPEPTRKWPWERTNTMNGSKSP 741

Search completed: September 20, 2005, 12:34:49  
Job time : 322.034 secs

Db 684 TEQKEDRNEDAEPIITAKAPSTSTPEPTRKWPWERTNTMNGSKSP 726  
|::|||::||: : : || | : : : || :  
RESULT 15  
AAW37153  
ID AAW37153 standard; protein; 802 AA.  
XX  
AC AAW37153;  
XX  
DT 06-JUL-1998 (first entry)  
XX  
DE Mouse neural Mena+++ protein.  
XX  
KW Neural Mena+++ protein; mammalian Ena; Enabled protein; Evi protein;  
KW cytoskeleton; cell morphology; cell adhesion; cell differentiation;  
KW cell growth; cell motility; mouse.  
XX  
OS Mus musculus.  
XX  
PN WO9801755-A1.  
XX  
PD 15-JAN-1998.  
XX  
PF 03-JUL-1997; 97MO-US011669.  
XX  
PR 05-JUL-1996; 96US-00675815.  
XX  
PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.  
PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.  
PI Gertler FB, Soriano P, Wehland J, Niebuhr K;  
XX  
DR WPI; 1998-101197/09.  
XX  
PT Detection of modulators of Mena and Ena-VASP-like genes and proteins -  
PT used in control of cytoskeletal dynamic events in normal and abnormal  
PT cell morphology, adhesion, motility, growth and differentiation.  
XX  
PS Example 4; Page 63-65; 77pp; English.  
XX  
CC This protein comprises novel murine neural Mena+.. Its amino acid  
CC sequence was deduced from a cDNA clone obtained from a mouse brain cDNA  
CC library. Two other isoforms, neural Mena+ (see AAW37151) and neural  
CC Mena++ (see AAW37152), are also disclosed. Unlike mammalian Ena (Mena,  
CC see AAW37148), neural Mena isoforms exhibit neural tissue-specific  
CC distribution. Based on the disclosed Mena and Evi genes (see also  
CC AAV02996-98) and proteins (see also AAW37148-49), a variety of methods  
CC and compositions are provided for screening, isolating and characterising  
CC endogenous and exogenous factors, drugs and therapeutic agents useful to  
CC evaluate and/or control cytoskeletal dynamic events involved in normal  
CC and abnormal cell morphology, adhesion, motility, growth and/or  
CC differentiation. A method of detecting a modulator of Mena  
CC activity/expression is claimed  
XX  
SQ Sequence 802 AA;

Query Match 11.8%; Score 336; DB 2; Length 802;  
Best Local Similarity 26.8%; Pred. No. 3.3e-12;  
Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;  
QY 4 PPALPP-PETAAATTAASSA---ASPHYEWILDTIDSLRSRKARPD----- 50  
DB 309 PTTSTPTTFLRHAATFATSLGSAFHVLPHYA-----TVPRPLNKNRSPSPVNTFSQ 363  
QY 51 --LERICRMVRRRHGPEPERTRAELE----KLQRAVLRSYKGSISYRNAARVQPPRR 104  
DB 364 PPAKSCAWTNSFSLPPPPPIIMISSPPCKATGPRPVLFCVSSPV-----PQM 413  
QY 105 GATTPAPPRA-----PRGAPAAAAAAPPPTPAPPPPPAPVAAAAAPAPRAA-----AA 154  
DB 414 PPSPTAPNGSLDSVTYVSPPTSGPAAAPPPPPPPPPPPPPPPPLPLPLASLSHCGS 473

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 81.6548 seconds  
(without alignments)  
491.841 Million cell updates/sec

Title: US-10-671-242-43  
Perfect score: 2850  
Sequence: 1 MAGPPALPPPTAAATAA.....HIKVLQGHFEDDDPGLG 538

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues  
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2850	100.0	538	4	US-09-616-289-43
2	2594	91.0	550	4	US-09-616-289-47
3	1445.5	50.7	317	3	US-08-979-608A-2
4	1445.5	50.7	317	4	US-09-517-849-2
5	1445.5	50.7	317	4	US-09-616-289-2
6	1191.5	41.8	252	3	US-08-979-608A-4
7	1191.5	41.8	252	4	US-09-517-849-4
8	1191.5	41.8	252	4	US-09-616-289-4
9	1170	41.1	217	3	US-08-979-608A-7
10	1170	41.1	217	4	US-09-517-849-7
11	1170	41.1	217	3	US-09-616-289-7
12	1086.5	38.1	232	3	US-08-979-608A-3
13	1086.5	38.1	232	4	US-09-517-849-3
14	1086.5	38.1	232	4	US-09-616-289-3
15	336	11.8	802	4	US-09-823-240A-2
16	319.5	11.2	783	4	US-09-252-991A-18035
17	301.5	10.6	8991	4	US-08-714-741-32
18	298.5	10.5	1958	1	US-07-945-283-2
19	298	10.5	581	4	US-09-949-016-9978
20	292.5	10.3	504	3	US-09-219-849-3
21	292.5	10.3	561	1	US-08-642-255-52
22	289.5	10.2	727	4	US-09-902-540-12383
23	288	10.1	330	1	US-08-642-255-32
24	288	10.1	408	1	US-07-609-716-65
25	288	10.1	408	3	US-08-475-411A-65
26	288	10.1	408	3	US-08-478-029A-65
27	286.5	10.1	960	3	US-09-219-849-5

28	286.5	10.1	1706	4	US-09-252-991A-31760	Sequence 31760, A
29	286.5	10.1	3122	4	US-10-237-551-201	Sequence 201, App
30	286.5	10.1	3122	4	US-10-237-551-250	Sequence 250, App
31	285.5	10.0	420	4	US-09-902-540-13993	Sequence 13993, A
32	284	10.0	357	1	US-07-609-716-66	Sequence 66, Appl
33	284	10.0	357	1	US-08-642-255-33	Sequence 33, Appl
34	284	10.0	357	3	US-08-475-411A-66	Sequence 66, Appl
35	284	10.0	357	3	US-08-478-029A-66	Sequence 4, Appl
36	282.5	9.9	720	3	US-09-219-849-4	Sequence 4, Appl
37	282.5	9.9	777	1	US-08-642-255-53	Sequence 53, Appl
38	281	9.9	905	2	US-08-574-959A-9	Sequence 9, Appl
39	281	9.9	905	3	US-09-357-014-9	Sequence 9, Appl
40	277.5	9.7	2142	4	US-09-538-092-1142	Sequence 1142, Ap
41	275.5	9.7	527	4	US-09-370-838-216	Sequence 216, App
42	275.5	9.7	527	4	US-09-854-133-216	Sequence 216, App
43	274.5	9.6	1135	2	US-08-574-959A-7	Sequence 7, Appl
44	274.5	9.6	1135	3	US-09-357-014-7	Sequence 7, Appl
45	269.5	9.5	726	4	US-09-252-991A-20675	Sequence 20675, A

ALIGNMENTS

RESULT 1  
US-09-616-289-43  
; Sequence 43, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-616-289-43

Query Match	100.0%;	Score	2850;	DB	4;	Length	538;
Best Local Similarity	100.0%;	Pred. No.	2.5e-172;				
Matches	538;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAGPPALPPPTAAATAA	SSAASPHYQEWILDTIDLSR	KARPDLEICRMVR	60		
Db	1	MAGPPALPPPTAAATAA	SSAASPHYQEWILDTIDLSR	KARPDLEICRMVR	60		
Qy	61	RHGPETRAEKLIOQRAVL	RVSKGISVRNAARVQPPRRGATPPAPR	PGAPA	120		
Db	61	RHGPETRAEKLIOQRAVL	RVSKGISVRNAARVQPPRRGATPPAPR	PGAPA	120		
Qy	121	AAAAAAPPPTAPP	PPAPVAAAAPARA	AAAAAATAPSPGPAQFPPRAQRAA	PLAAP	180	
Db	121	AAAAAAPPPTAPP	PPAPVAAAAPARA	AAAAAATAPSPGPAQFPPRAQRAA	PLAAP	180	
Qy	181	PPAPAPPAVAPPAG	RRRPPPAVAREP	LPPLPPPPAPPQQQPPPPPPGGAV	240		
Db	181	PPAPAPPAVAPPAG	RRRPPPAVAREP	LPPLPPPPAPPQQQPPPPPPGGAV	240		

```
QY 241 RAGGAARPVSLREVYVYLGSGGAGGLTRGRVQGLLEBBAARGRLERTRIGALALPRG 300
Db 241 RAGGAARPVSLREVYVYLGSGGAGGLTRGRVQGLLEBBAARGRLERTRIGALALPRG 300
QY 301 DRPGRAPPASARPSRKRGGEERVLKEKEEEDDEDEDDVSGSEVPESDRPAGAQ 360
Db 301 DRPGRAPPASARPSRKRGGEERVLKEKEEEDDEDEDDVSGSEVPESDRPAGAQ 360
QY 361 HHQNGERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFSMAANKKEGGTASVA 420
Db 361 HHQNGERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFSMAANKKEGGTASVA 420
QY 421 TGPDSFSPVPLPGKAPALPGADGTPGCPGPRKEKSPDVVMTVMDVVEYFTAGPPEQA 480
Db 421 TGPDSFSPVPLPGKAPALPGADGTPGCPGPRKEKSPDVVMTVMDVVEYFTAGPPEQA 480
QY 481 TAFQEQEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCGHFEDDDPDGFLG 538
Db 481 TAFQEQEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCGHFEDDDPDGFLG 538

RESULT 2
US-09-616-289-47
; Sequence 47, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-47

Query Match 91.0%; Score 2594; DB 4; Length 550;
Best Local Similarity 91.1%; Pred. No. 3.6e-156;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPTAAAAATT-AAAASSSAASPHYQEWILDTIDSLSRKRKARPDLEICRMVR 59
Db 1 MAGPPALPPPTAAAAATTAAAAASSSAASPHYQEWILDTIDSLSRKRKARPDLEICRMVR 60
QY 60 RRGGPPERTRALEKLIQRAVLRYVSYKGSISYRNAARVQPPRRGATPPAPPRARGAP 119
Db 61 RRGGPPERTRALEKLIQRAVLRYVSYKGSISYRNAARVQPPRRGATPPAPPRARGGP 120
QY 120 AAAAAAAPPPTAPPPPPAPV-AAAAAPARAPR-AAAAAATAPSPGPAQPGPRAORAP 176
Db 121 --AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAAATAPSPGPAQPGPRAORAP 178
QY 177 LAAPPAPAPAPVAPAGPRRAPP--AVAAREPLPPPPPPPPPPPPPPPPPPPPPPPP 229
Db 179 LAAPPAPAPAPVAPAGPRRAPPAAAAVAARESPLPPPPPPPPPPPPPPPPPPPPPP 238
```

```
QY 230 QPOPPEGCAGVAGGAARPVSLREVYVYLGSGGAGGLTRGRVQGLLEBBAARGRLER 289
Db 239 QPOPPEGCAGVAGGAARPVSLREVYVYLGSGGAGGLTRGRVQGLLEBBAARGRLER 298
QY 290 TRIGALALPRGDRPGRAPPASARPSRKRGGEERVLKEKEEEDDEDEDDVSGSEVP 347
Db 299 TRIGALALPRGDRPGRAPPASARPSRKRGGEERVLKEKEEEDDEDEDDVSGSEVP 358
QY 348 SEVPESDRPAGAHQHLN--GERGQSAKERVKEWTPCGPHQGDGGRGAPGSGTRQVFS 406
Db 359 SEVPESDRPAGAHQHLNGERGPOTAKERAKEWSLCGPHPGQEGRGGAAGSGTRQVFS 418
QY 407 MAAMNKEGTSASVATGPDSPSPVPLPGKAPALPGADGTPGCPGPRKEKSPDVVMTVMD 466
Db 419 MAALSKEGGSASTTGPDSFSPVPLPGKAPALPGADGTPGCPGPRKEKSPDVVMTVMD 478
QY 467 VVEYFTAGFPQATAFQEQEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCG 526
Db 479 VVEYFTAGFPQATAFQEQEIDGKSLLMQRTDVLGLSIRLGPALKIYEHHIKVLQCG 538
QY 527 HFEDDDPDGFLG 538
Db 539 HFEDDDPDGFLG 550

RESULT 3
US-08-979-608A-2
; Sequence 2, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NO. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-979-608A-2
```



```
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-517-849-2

Query Match      50.7%; Score 1445.5; DB 3; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-84;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 294
Db 11 PKGAARAGGAPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 70
QY 295 LALPRGDRPGRAPPAASARPSRKSGEERVLKEKEEEDDEDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARARNRKAGEERVLKEKEEEDDEDDDDVVSEGSEVPE 130
QY 353 SDRPAGAQQHQLN--GERGPSAKERVKWTPCGPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQQHQLNGGGRGPQTKERAKESLCPGHPGQEGRGAPAGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGPDSPSPVPLPGKPKALPGADGTPPGCPGRKEKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTTGPDPSPSPVPLPGKPKALPGADGTPPGCPGRKEKPADPVETVMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHKIKVLOQGHFEDD 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHKIKVLOQGHFEDD 310
QY 532 DPGFLG 538
Db 311 DPEGFLG 317

RESULT 4
US-09-517-849-2
; Sequence 2, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-517-849-2

Query Match      50.7%; Score 1445.5; DB 4; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-84;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 294
Db 11 PKGAARAGGAPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 70
QY 295 LALPRGDRPGRAPPAASARPSRKSGEERVLKEKEEEDDEDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARARNRKAGEERVLKEKEEEDDEDDDDVVSEGSEVPE 130
QY 353 SDRPAGAQQHQLN--GERGPSAKERVKWTPCGPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQQHQLNGGGRGPQTKERAKESLCPGHPGQEGRGAPAGSGTRQVFSMAALS 190
QY 412 KEGGTASVATGPDSPSPVPLPGKPKALPGADGTPPGCPGRKEKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTTGPDPSPSPVPLPGKPKALPGADGTPPGCPGRKEKPADPVETVMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHKIKVLOQGHFEDD 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLPALKIYEHKIKVLOQGHFEDD 310
QY 532 DPGFLG 538
Db 311 DPEGFLG 317

RESULT 5
US-09-616-289-2
; Sequence 2, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-616-289-2

Query Match      50.7%; Score 1445.5; DB 4; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-84;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGA 294
```

Db 11 PKGAARAGAPVSLREVRVYLGSSGAGRLTRGRVQGLLEBAARGLRTRLGA 70  
QY 295 LALPRGDRPGRAPPAASARPSKRKGEEERVLKEKEEEDDDDEEDD--VSEGSEVPE 352  
Db 71 LALPRGDRPGRAPPAASARAARNAKRAGEERVLKEKEEEDDDDDVVSSEGSEVPE 130  
QY 353 SDPAGAAHOLN-GERGPOSAKERVKWTPCGPHQOQDEGRGPAPGSGTROVFSMAANV 411  
Db 131 SDPAGAAHQLNGGREGPQTAKERAKEWSLCGPHPGQEGRGPAAGSGTROVFSMAALS 190  
QY 412 KEGGTASVATGPDSPSPVPLPPGKPALPGADGTPFCGPKRKEKPSDPVEMTMDVVEYF 471  
Db 191 KEGGSASSTGPDSPSPVPLPPGKPALPGADGTPFCGPKRKEKPADPVEMTMDVVEYF 250  
QY 472 TRAGFPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVQQGHFEDD 531  
Db 251 TRAGFPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVQQGHFEDD 310  
QY 532 DPGDPLG 538  
Db 311 DPEGFLG 317

## RESULT 6

US-08-979-608A-4  
; Sequence 4, Application US/08979608A  
; Patent No. 6355451  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/979,608A  
; FILING DATE: 26-No. 6355451-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/048,547  
; FILING DATE: 03-JUN-1997  
; APPLICATION NUMBER: US 60/031,930  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 252 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-979-608A-4

Query Match 41.8%; Score 1191.5; DB 3; Length 252;  
Best Local Similarity 88.5%; Pred. No. 4.6e-68;  
Matches 223; Conservative 17; Mismatches 9; Indels 3; Gaps 2;  
QY 290 TRLGALALPRGDRPGRAPPAASARPSKRKGEEERVLKEKEEEDDDDEEDD--VSEG 347  
Db 1 TRLGALALPRGDRPGRAPPAASARAARNAKRAGEERVLKEKEEEDDDDDVVSSEG 60  
QY 348 SVPSPSDRPAGAAHOLN-GERGPOSAKERVKWTPCGPHQOQDEGRGPAPGSGTROVFS 406  
Db 61 SVPSPSDRPAGAAHQLNGGREGPQTAKERAKEWSLCGPHPGQEGRGPAAGSGTROVFS 120  
QY 407 MAAMNKEGTTASVATGPDSPSPVPLPPGKPALPGADGTPFCGPKRKEKPSDPVEMTMD 466  
Db 121 MAALSKEGGSASSTGPDSPSPVPLPPGKPALPGADGTPFCGPKRKEKPADPVEMTMD 180  
QY 467 VVEYFTEAGFPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVQQ 526  
Db 181 VVEYFTEAGFPEQATAFOEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVQQ 240  
QY 527 HFEDDDPDGFLG 538  
Db 241 HFEDDDPDGFLG 252

## RESULT 7

US-09-517-849-4  
; Sequence 4, Application US/09517849  
; Patent No. 6605588  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/517,849  
; FILING DATE: 02-Mar-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/979,608  
; FILING DATE: 26-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 252 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-517-849-4  
Query Match 41.8%; Score 1191.5; DB 4; Length 252;

[illegible]

Db 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217  
|||||

RESULT 10

US-09-517-849-7  
; Sequence 7, Application US/09517849  
; Patent No. 6605588  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/517,849  
; FILING DATE: 02-Mar-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/979,608  
; FILING DATE: 26-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-517-849-7

Query Match 41.1%; Score 1170; DB 4; Length 217;  
Best Local Similarity 100.0%; Pred. No. 8.9e-67;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 322 EERVLEKEEEEDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 381  
Db 1 EERVLEKEEEEDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 60  
QY 382 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPCKPALPGA 441  
Db 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPCKPALPGA 120  
QY 442 DGTFFGCGPGRKEKPSDPVWTVMDVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 501  
Db 121 DGTFFGCGPGRKEKPSDPVWTVMDVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 180  
QY 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538  
Db 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217

RESULT 11

US-09-616-289-7  
; Sequence 7, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/616,289  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-616-289-7  
Query Match 41.1%; Score 1170; DB 4; Length 217;  
Best Local Similarity 100.0%; Pred. No. 8.9e-67;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 322 EERVLEKEEEEDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 381  
Db 1 EERVLEKEEEEDDEDEDDVSEGVESVPSDRPAGAHQHLNGERGQSAKERVKEWT 60  
QY 382 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPCKPALPGA 441  
Db 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPCKPALPGA 120  
QY 442 DGTFFGCGPGRKEKPSDPVWTVMDVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 501  
Db 121 DGTFFGCGPGRKEKPSDPVWTVMDVVEYFTAGPPEQATAFQEQIDGKSLLLMORTDV 180  
QY 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538  
Db 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217  
RESULT 12  
US-08-979-608A-3  
; Sequence 3, Application US/08979608A  
; Patent No. 6355451  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette

```
/
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/979,608A
/ FILING DATE: 26-NOV-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/048,547
/ FILING DATE: 03-JUN-1997
/ APPLICATION NUMBER: US 60/031,930
/ FILING DATE: 27-NOV-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 232 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match      38.1%; Score 1086.5; DB 3; Length 232;
Best Local Similarity 87.5%; Pred. No. 1.8e-61;
Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;

QY 310 ASARPSRKGGREVLKEEEDDEDEDEDD--VSEGSEVPESDRPAGAHQHLN-G 366
Db 1 ASARAARNKGAEEERVLKEEEDDEDEDEDDDDVVSEGSEVPESDRPAGAHQHLNCG 60

QY 367 ERGPQAKERVKWTCPGPHQGDGREGPAGSGTRQVFSMAANKEGGTASVATGPDSP 426
Db 61 ERGPQAKERAKESWLCGPHGQEGREGPAAGSGTRQVFSMAALSKEGGSASSTTGPDS 120

QY 427 SPVPLPGKALPGADTTPGCPGPKGKPSDPVMTVMDVYFTFTEAGPPEQATAFQEQ 486
Db 121 SPVPLPGKALPGADTTPGCPGPKGKPADPVMTVMDVYFTFTEAGPPEQATAFQEQ 180

QY 487 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 538
Db 181 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 232

RESULT 13
US-09-517-849-3
; Sequence 3, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
```

```
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/517,849
/ FILING DATE: 02-Mar-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/979,608
/ FILING DATE: 26-NOV-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-003001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 232 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-517-849-3

Query Match      38.1%; Score 1086.5; DB 4; Length 232;
Best Local Similarity 87.5%; Pred. No. 1.8e-61;
Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;

QY 310 ASARPSRKGGREVLKEEEDDEDEDEDD--VSEGSEVPESDRPAGAHQHLN-G 366
Db 1 ASARAARNKGAEEERVLKEEEDDEDEDEDDDDVVSEGSEVPESDRPAGAHQHLNCG 60

QY 367 ERGPQAKERVKWTCPGPHQGDGREGPAGSGTRQVFSMAANKEGGTASVATGPDSP 426
Db 61 ERGPQAKERAKESWLCGPHGQEGREGPAAGSGTRQVFSMAALSKEGGSASSTTGPDS 120

QY 427 SPVPLPGKALPGADTTPGCPGPKGKPSDPVMTVMDVYFTFTEAGPPEQATAFQEQ 486
Db 121 SPVPLPGKALPGADTTPGCPGPKGKPADPVMTVMDVYFTFTEAGPPEQATAFQEQ 180

QY 487 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 538
Db 181 EIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQGHFEDDDPDGFLG 232

RESULT 14
US-09-616-289-3
; Sequence 3, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-3
```

Query Match 38.1%; Score 1086.5; DB 4; Length 232;  
Best Local Similarity 87.5%; Pred. No. 1.8e-61;  
Matches 203; Conservative 17; Mismatches 9; Indels 3; Gaps 2;  
QY 310 ASARPSRKRGGSERVLEKEEREDDDEDEED--VSEGEVPESDRPAQAQHOLN-G 366  
Db 1 ASARAARNKRAGSERVLEKEEREDDDEDEEDVSEGEVPESDRPAQAQHOLNGG 60  
QY 367 ERGPQSAKERVKEWTTCGPHQGDGEGRPAGSGTRQVFSMAANKEGGTASVATGPDSP 426  
Db 61 ERGPQAKERAKESLSCGPHPGQEGRGPAAGSGTRQVFSMAALSKEGGSASSTTCGPDSP 120  
QY 427 SPVPLPGKALPGAGTTPGCGPKGKEKPSDPEVMTVMDVVEYFTEAGPPEQATAFQEQ 486  
Db 121 SPVPLPGKALPGAGTTPGCGPKGKEKPADPEVMTVMDVVEYFTEAGPPEQATAFQEQ 180  
QY 487 EIDGKSILLMORTDVLTLGSLRGLPALKIYEHHIKVLQOQHFDDEDDPDGFLG 538  
Db 181 EIDGKSILLMORTDVLTLGSLRGLPALKIYEHHIKVLQOQHFDDEDDPDGFLG 232

Search completed: September 20, 2005, 12:45:32  
Job time : 82.6548 secs

RESULT 15  
US-09-823-240A-2  
; Sequence 2, Application US/09823240A  
; Patent No. 6716597  
; GENERAL INFORMATION:  
; APPLICANT: Frank B. Gertler  
; APPLICANT: James E. Bear  
; APPLICANT: Jurgen Wehland  
; APPLICANT: Joseph Loureio  
; TITLE OF INVENTION: Methods and Products for Regulating Cell  
; TITLE OF INVENTION: Motility  
; FILE REFERENCE: M00656.70064.US  
; CURRENT APPLICATION NUMBER: US/09/823,240A  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/194,564  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 802  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-823-240A-2

Query Match 11.8%; Score 336; DB 4; Length 802;  
Best Local Similarity 26.8%; Pred. No. 1.4e-13;  
Matches 124; Conservative 44; Mismatches 169; Indels 126; Gaps 21;  
QY 4 PPALPP-PETAAATTAASSA---ASPHYQEWILDTIDLSRSRKARP-----50  
Db 309 PTSTPTPLRHAATRFATSLGSAFHPVLPHYA----TVPRPLNKNRSPSPVNTPSQ 363  
QY 51 --LERICRMVRRHGPETRAELE---KLQORAVLRVSYKGSISYRNAARVQPPRR 104  
Db 364 PPAKSCAMPTSNFSLPSPSPIMISSPGKATGPRVLPVCVSSPV-----POM 413  
QY 105 GATPPAPPRA-----PRGAPAAAAAAPPPTAPPPAPVAAAAAPAPRAAA---AA 154  
Db 414 PPSPTAPNGSLDVTYVSPPTSGPAAPPPPPPPPPPPPPPPPPPPPPPLPLPLASLSHCGS 473  
QY 155 ANTAPPS-----PGP-----AOPGPRQRAAPLAA 179  
Db 474 QASFPPTPLASTPSSKPSVLPSPSAGAPASAEPLNPELGDSASEPGLQA-ASQPAES 532  
QY 180 P-----PPAAPP-----AVAPAGPRRAPPAVAAREPPLPPPPQPPAPP 221  
Db 533 PTPQGLVLPAPPPPPPLPSGPAYASALPPPPGP--PPPPPLPSTGPPPPPPPPPPLEN 590  
QY 222 QQQQPPPPQPPPEGGAVRAGAA---RPVS-----LRVYVYLGSGGAGGRL 268  
Db 591 QAPPPPPPPAPPLPASGIFSGTSDNRLPTGLAAAIAGAKLRKYSRVEDGSGFFPGG-- 648

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 296.987 Seconds  
(without alignments)  
733.538 Million cell updates/sec

Title: US-10-671-242-43

Perfect score: 2850

Sequence: 1 MAGPPALPPPTAAATTA.....HIKVLQGHFDDPDGFLG 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2850	100.0	538	9	US-09-976-740-43
2	2850	100.0	538	13	US-10-023-529-43
3	2850	100.0	538	15	US-10-023-523-43
4	2850	100.0	538	15	US-10-023-523-43
5	2850	100.0	538	15	US-10-023-523-43
6	2594	91.0	550	9	US-09-976-740-47
7	2594	91.0	550	13	US-10-023-529-47
8	2594	91.0	550	13	US-10-023-523-47
9	2594	91.0	550	15	US-10-023-523-47
10	2594	91.0	550	15	US-10-023-523-47
11	1445.5	50.7	317	9	US-09-962-055-2

12	1445.5	50.7	317	9	US-09-976-740-2	Sequence 2, Appli
13	1445.5	50.7	317	13	US-10-023-529-2	Sequence 2, Appli
14	1445.5	50.7	317	13	US-10-023-523-2	Sequence 2, Appli
15	1445.5	50.7	317	15	US-10-023-523-2	Sequence 2, Appli
16	1445.5	50.7	317	15	US-10-023-523-2	Sequence 2, Appli
17	1260	44.2	241	9	US-09-925-298-665	Sequence 665, App
18	1260	44.2	241	14	US-10-103-806-665	Sequence 665, App
19	1191.5	41.8	252	9	US-09-962-055-4	Sequence 4, Appli
20	1191.5	41.8	252	9	US-09-962-055-4	Sequence 4, Appli
21	1191.5	41.8	252	13	US-10-023-529-4	Sequence 4, Appli
22	1191.5	41.8	252	13	US-10-023-529-4	Sequence 4, Appli
23	1191.5	41.8	252	15	US-10-023-523-4	Sequence 4, Appli
24	1191.5	41.8	252	15	US-10-023-523-4	Sequence 4, Appli
25	1170	41.1	217	9	US-09-962-055-7	Sequence 7, Appli
26	1170	41.1	217	9	US-09-962-055-7	Sequence 7, Appli
27	1170	41.1	217	13	US-10-023-529-7	Sequence 7, Appli
28	1170	41.1	217	13	US-10-023-529-7	Sequence 7, Appli
29	1170	41.1	217	15	US-10-023-523-7	Sequence 7, Appli
30	1170	41.1	217	15	US-10-023-523-7	Sequence 7, Appli
31	1086.5	38.1	232	9	US-09-962-055-3	Sequence 3, Appli
32	1086.5	38.1	232	9	US-09-962-055-3	Sequence 3, Appli
33	1086.5	38.1	232	13	US-10-023-529-3	Sequence 3, Appli
34	1086.5	38.1	232	13	US-10-023-529-3	Sequence 3, Appli
35	1086.5	38.1	232	15	US-10-023-523-3	Sequence 3, Appli
36	1086.5	38.1	232	15	US-10-023-523-3	Sequence 3, Appli
37	367	12.9	321	16	US-10-425-115-202316	Sequence 202316,
38	361	12.7	431	16	US-10-437-963-204963	Sequence 204963,
39	342.5	12.0	376	16	US-10-437-963-154970	Sequence 154970,
40	340.5	11.9	270	16	US-10-437-963-130068	Sequence 130068,
41	340	11.9	215	16	US-10-767-701-35371	Sequence 35371, A
42	336	11.8	296	16	US-10-437-963-136002	Sequence 136002,
43	336	11.8	802	9	US-09-823-240-2	Sequence 2, Appli
44	335.5	11.8	252	16	US-10-437-963-151853	Sequence 151853,
45	335.5	11.8	409	16	US-10-437-963-140753	Sequence 140753,

#### ALIGNMENTS

#### RESULT 1

US-09-976-740-43  
; Sequence 43, Application US/09976740  
; Publication No. US20020194633A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/976,740  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 538  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-740-43

Query Match 100.0%; Score 2850; DB 9; Length 538;  
Best Local Similarity 100.0%; Pred. No. 4e-131;  
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPPALPPETAAATAAASSAASHYQEWILDTIDSLRSRKARPDLERICMVR 60  
DB 1 MAGPPALPPETAAATAAASSAASHYQEWILDTIDSLRSRKARPDLERICMVR 60  
QY 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
QY 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180  
DB 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180  
QY 181 PPAPAAPPAVAPAGPRRRAPPAAREPPLPPPOPPAPPOQOQPPPOPPPPGGAV 240  
DB 181 PPAPAAPPAVAPAGPRRRAPPAAREPPLPPPOPPAPPOQOQPPPOPPPPGGAV 240  
QY 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300  
DB 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300  
QY 301 DRPGRAPPAASARPSRKSGEERVLEKEEEDDDDEEDDDVSGSEVPESDRPAGAQ 360  
DB 301 DRPGRAPPAASARPSRKSGEERVLEKEEEDDDDEEDDDVSGSEVPESDRPAGAQ 360  
QY 361 HHQNGERGQSAKERVKETPCGPHQODEGRGAPGSGTRQVFSMAANKEGGTASVA 420  
DB 361 HHQNGERGQSAKERVKETPCGPHQODEGRGAPGSGTRQVFSMAANKEGGTASVA 420  
QY 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKRKEKPSDPVETVMDVVEYFTEAGPPEQA 480  
DB 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKRKEKPSDPVETVMDVVEYFTEAGPPEQA 480  
QY 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538  
DB 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538

RESULT 2  
US-10-023-529-43  
; Sequence 43, Application US/10023529  
; Publication No. US20020129388A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/10/023,529  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 538  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-023-529-43

Query Match 100.0%; Score 2850; DB 13; Length 538;  
Best Local Similarity 100.0%; Pred. No. 4e-131;

Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAGPPALPPETAAATAAASSAASHYQEWILDTIDSLRSRKARPDLERICMVR 60  
DB 1 MAGPPALPPETAAATAAASSAASHYQEWILDTIDSLRSRKARPDLERICMVR 60  
QY 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
DB 61 RHGPEPERTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAGAPA 120  
QY 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180  
DB 121 AAAAAAPPPTAPPPPPAPVAAAAPARAPRAAAAAATAPSPGPAQPGPRAQRAAPLAAP 180  
QY 181 PPAPAAPPAVAPAGPRRRAPPAAREPPLPPPOPPAPPOQOQPPPOPPPPGGAV 240  
DB 181 PPAPAAPPAVAPAGPRRRAPPAAREPPLPPPOPPAPPOQOQPPPOPPPPGGAV 240  
QY 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300  
DB 241 RAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLRLGALALPRG 300  
QY 301 DRPGRAPPAASARPSRKSGEERVLEKEEEDDDDEEDDDVSGSEVPESDRPAGAQ 360  
DB 301 DRPGRAPPAASARPSRKSGEERVLEKEEEDDDDEEDDDVSGSEVPESDRPAGAQ 360  
QY 361 HHQNGERGQSAKERVKETPCGPHQODEGRGAPGSGTRQVFSMAANKEGGTASVA 420  
DB 361 HHQNGERGQSAKERVKETPCGPHQODEGRGAPGSGTRQVFSMAANKEGGTASVA 420  
QY 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKRKEKPSDPVETVMDVVEYFTEAGPPEQA 480  
DB 421 TGPDSPSVPLPPGKALPGADGTPGCPGPKRKEKPSDPVETVMDVVEYFTEAGPPEQA 480  
QY 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538  
DB 481 TAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQGHFEDDDPDGFLG 538

RESULT 3  
US-10-023-523-43  
; Sequence 43, Application US/10023523  
; Publication No. US20020152485A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/10/023,523  
; CURRENT FILING DATE: 2001-12-17  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 538  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-023-523-43

Query Match 100.0%; Score 2850; DB 13; Length 538;



Best Local Similarity 100.0%, Pred. No. 4e-131; Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAGPPALPPPTATAAAATATAAASSAASPHYQEWILDTIDSLSRKARPDLERICRMVRR	60
Db	1	MAGPPALPPPTATAAAATATAAASSAASPHYQEWILDTIDSLSRKARPDLERICRMVRR	60
Qy	61	RHGPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRGATPPAPPRARGAPA	120
Db	61	RHGPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRGATPPAPPRARGAPA	120
Qy	121	AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP	180
Db	121	AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP	180
Qy	181	PPAPAAAPVAVAPPGRRAPPPVAVAAAREPPLPPPPQPPAPPPQOQQPPPPQPPPEGAV	240
Db	181	PPAPAAAPVAVAPPGRRAPPPVAVAAAREPPLPPPPQPPAPPPQOQQPPPPQPPPEGAV	240
Qy	241	RAGGAARPVSUREVRYLGGSGGAGRLTRGRVQGLLEEEAAAAGRLETRLGALALPRG	300
Db	241	RAGGAARPVSUREVRYLGGSGGAGRLTRGRVQGLLEEEAAAAGRLETRLGALALPRG	300
Qy	301	DRGRRAPPAASARPSRSKRGEERVLKEKEEEDDEDEDEDDVSEGSVEPESDRPAGAQ	360
Db	301	DRGRRAPPAASARPSRSKRGEERVLKEKEEEDDEDEDEDDVSEGSVEPESDRPAGAQ	360
Qy	361	HHQLNGERGPOSASAKERVKEWTPCCGPHQOQDEGRGPAQSGGTRQVFSMAAMNKEGGTASVA	420
Db	361	HHQLNGERGPOSASAKERVKEWTPCCGPHQOQDEGRGPAQSGGTRQVFSMAAMNKEGGTASVA	420
Qy	421	TGPDSPSPVPLPPGKPALPGADGTFPGCPGCKPKPSDPVSWTVMDDVVVEYTEAGFPEQA	480
Db	421	TGPDSPSPVPLPPGKPALPGADGTFPGCPGCKPKPSDPVSWTVMDDVVVEYTEAGFPEQA	480
Qy	481	TAFOEQSIDGKSLLLMORTDVLTLGLSIRLGPAIKIYEHHIKVLOOQHFDDEDDPDGFLG	538
Db	481	TAFOEQSIDGKSLLLMORTDVLTLGLSIRLGPAIKIYEHHIKVLOOQHFDDEDDPDGFLG	538

Query Match	100.0%	Score 2850;	DB 15;	Length 538;
Beat Local Similarity	100.0%;	Pred. No. 4e-131;		
Matches 538;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAGPPALPPPTTAAATATAAASSSAASPHYQEWILDTISLSRSRKARPDLERICMVR	60	
Db	1	MAGPPALPPPTTAAATATAAASSSAASPHYQEWILDTISLSRSRKARPDLERICMVR	60	
Qy	61	RHGPEPERTRAELEKLIQORAVLSVYKGSISYRNAARVQPPRRGATPPAPPRGAPA	120	
Db	61	RHGPEPERTRAELEKLIQORAVLSVYKGSISYRNAARVQPPRRGATPPAPPRGAPA	120	
Qy	121	AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP	180	
Db	121	AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAATAPPSPGPAQPGPRAQRAAPLAAP	180	
Qy	181	PPAAPAAPVAPPAGPRAPPAPVAAAREPPLPPPPPPAPQOQPPPPQPPPEGGAV	240	
Db	181	PPAAPAAPVAPPAGPRAPPAPVAAAREPPLPPPPPPAPQOQPPPPQPPPEGGAV	240	
Qy	241	RAGGAARFVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAAARGRLTRLGALALPRG	300	
Db	241	RAGGAARFVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAAARGRLTRLGALALPRG	300	
Qy	301	DRGRAPPAASARPSRKRGGEERVLKKEEEDDEDEDDVSSEGESEVPSDRPAGAQ	360	
Db	301	DRGRAPPAASARPSRKRGGEERVLKKEEEDDEDEDDVSSEGESEVPSDRPAGAQ	360	
Qy	361	HHQLNGRGQSASAKERVKWETPCPHQODEGRGPAFGSGTRQVFSMAANKEGGTASVA	420	
Db	361	HHQLNGRGQSASAKERVKWETPCPHQQDGRGPAFGSGTRQVFSMAANKEGGTASVA	420	
Qy	421	TGPDSPSPVPLPCKPALPGADGTPFCGPPGRKEKPSDPVEMTWMDVVVEYFTEAGFPBQA	480	
Db	421	TGPDSPSPVPLPCKPALPGADGTPFCGPPGRKEKPSDPVEMTWMDVVVEYFTEAGFPBQA	480	
Qy	481	TAFOEQIDGKSLLLMORTDVLITGLSIRLGPALKIYEHHIKVLQOGHFDDDDPDGFLG	538	
Db	481	TAFOEQIDGKSLLLMORTDVLITGLSIRLGPALKIYEHHIKVLQOGHFDDDDPDGFLG	538	



```
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-529-47

Query Match      91.0%; Score 2594; DB 13; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPETAAAAATT-AAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRRCMVR 59
Db 1 MAGPPALPPPETAAAAATTAAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRRCMVR 60
QY 60 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPRPRGAP 119
Db 61 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPRPRGAP 120
QY 120 AAAAAAAPPPTAPPPPPAPV-AAAAPARAPR--AAAAAATAPPSGPAQPGPRAQAAP 176
Db 121 --AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAAATAAPPSGPAQPGPRAQAAP 178
QY 177 LAAPPAPAPAPVAPAGPRRAPPP--AVAAREPPLPPPPQPPAPPPQOOQ-----PPPP 229
Db 179 LAAPPAPAPAPVAPAGPRRAPPPAAVAARESPPLPPPPQPPAPPPQOOQPPPPPPQ 238
QY 230 QOQPPEGGAVRAGGAARPVSLREVRYLGGSGGAGRLTRGRVQGLLEEEAAARGRLER 289
Db 239 QOQPPEGGAAAGGAPRVPVSLREVRYLGGSSGAGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAPPAASARPSKGGSERVLEKEEEDDDDEDEEDD--VSEG 347
Db 299 TRIGALALPRGDRPGRAPPAASARAARNKRAAGEERVLEKEEEDDDDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN--GERGQSAKERVKEWTPCGPHOGODEGRGPAKSGSTRQVFS 406
Db 359 SEVPESDRPAGAQHQLNGERGPQTAKERAKESWLCGPHGQEGERGPAKSGSTRQVFS 418
QY 407 MAAMNKEGGTASVATGPDSPSPVPLPGKPALPGADGTFPGCPGRKEKPSDPVETVMD 466
Db 419 MAALSKEGGSASSTGTPDPSFSPVPLPGKPALPGADGTFPGCPGRKEKPADPVENTVMD 478
QY 467 VVEYFTEAGFPQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 526
Db 479 VVEYFTEAGFPQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 538
QY 527 HFEDDDPDGFLG 538
Db 539 HFEDDDPEGFLG 550

RESULT 8
US-10-023-523-47
; Sequence 47, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27

; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-523-47

Query Match      91.0%; Score 2594; DB 13; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPPETAAAAATT-AAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRRCMVR 59
Db 1 MAGPPALPPPETAAAAATTAAAASSSAASPHYQEWILDTIDSLSRSKARPDLSRRCMVR 60
QY 60 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPRPRGAP 119
Db 61 RRHGPEPERTRAELEKLIQORAVLRSYKGSISYRNAARVQPPRRGATPPAPRPRGAP 120
QY 120 AAAAAAAPPPTAPPPPPAPV-AAAAPARAPR--AAAAAATAPPSGPAQPGPRAQAAP 176
Db 121 --AAAAAAPPPTAPPPPPAPVAAAAAPARAPRAAAAAAATAAPPSGPAQPGPRAQAAP 178
QY 177 LAAPPAPAPAPVAPAGPRRAPPP--AVAAREPPLPPPPQPPAPPPQOOQ-----PPPP 229
Db 179 LAAPPAPAPAPVAPAGPRRAPPPAAVAARESPPLPPPPQPPAPPPQOOQPPPPPPQ 238
QY 230 QOQPPEGGAVRAGGAARPVSLREVRYLGGSGGAGRLTRGRVQGLLEEEAAARGRLER 289
Db 239 QOQPPEGGAAAGGAPRVPVSLREVRYLGGSSGAGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAPPAASARPSKGGSERVLEKEEEDDDDEDEEDD--VSEG 347
Db 299 TRIGALALPRGDRPGRAPPAASARAARNKRAAGEERVLEKEEEDDDDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN--GERGQSAKERVKEWTPCGPHOGODEGRGPAKSGSTRQVFS 406
Db 359 SEVPESDRPAGAQHQLNGERGPQTAKERAKESWLCGPHGQEGERGPAKSGSTRQVFS 418
QY 407 MAAMNKEGGTASVATGPDSPSPVPLPGKPALPGADGTFPGCPGRKEKPSDPVETVMD 466
Db 419 MAALSKEGGSASSTGTPDPSFSPVPLPGKPALPGADGTFPGCPGRKEKPADPVENTVMD 478
QY 467 VVEYFTEAGFPQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 526
Db 479 VVEYFTEAGFPQATAFQOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQOG 538
QY 527 HFEDDDPDGFLG 538
Db 539 HFEDDDPEGFLG 550

RESULT 9
US-10-616-187-47
; Sequence 47, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
```

```
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-616-187-47

Query Match      91.0%; Score 2594; DB 15; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPETAAATT-AAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 59
DB 1 MAGPPALPPETAAATTAAAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
QY 60 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 119
DB 61 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 120
QY 120 AAAAAAAPTTPAPPPAPV-AAAAAPARPR--AAAAAATAPPSGPAQPGPRAQRAAP 176
DB 121 --AAAAAPTTPAPPPAPVAAAAAPARPRAAAAAATAATAPPSGPAQPGPRAQRAAP 178
QY 177 LAAPPAPAAAPAVAPAGPRAPP--AFAAREPPLPPPPQPPQOQO-----PPPP 229
DB 179 LAAPPAPAAAPAVAPAGPRAPPAAAAAARESPPLPPPPQPPQOQOQPPPPPPQ 238
QY 230 QOPPPPEGAVRAGGAARPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 289
DB 239 QOPPPPEGGAARAGGARPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAAPAAASARPSKRGGEERVLEKEEEDDEDEDD--VSEG 347
DB 299 TRIGALALPRGDRPGRAAPAAASARAARKRAGEERVLEKEEEDDEDEDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN-GERGQSAKERVKEWTPCGPHQODEGRGAPGSGTRQVFS 406
DB 419 MAALSKEGGSASTTGDSPSPVPLPPGKPALPGADGTPFGCPAGRKEKPADPVEVTMD 478
QY 467 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHHIKVLOOG 526
DB 479 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHHIKVLOOG 538
QY 527 HFEDDDPDGFLG 538
DB 539 HFEDDDPEGFLG 550

RESULT 10
US-10-671-242-47
; Sequence 47, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001

; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-671-242-47

Query Match      91.0%; Score 2594; DB 15; Length 550;
Best Local Similarity 91.1%; Pred. No. 1.2e-118;
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;

QY 1 MAGPPALPPETAAATT-AAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 59
DB 1 MAGPPALPPETAAATTAAAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVR 60
QY 60 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 119
DB 61 RRHGPPEPTRAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPPRGAP 120
QY 120 AAAAAAAPTTPAPPPAPV-AAAAAPARPR--AAAAAATAPPSGPAQPGPRAQRAAP 176
DB 121 --AAAAAPTTPAPPPAPVAAAAAPARPRAAAAAATAATAPPSGPAQPGPRAQRAAP 178
QY 177 LAAPPAPAAAPAVAPAGPRAPP--AFAAREPPLPPPPQPPQOQO-----PPPP 229
DB 179 LAAPPAPAAAPAVAPAGPRAPPAAAAAARESPPLPPPPQPPQOQOQPPPPPPQ 238
QY 230 QOPPPPEGAVRAGGAARPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 289
DB 239 QOPPPPEGGAARAGGARPVSLREVVRVYLGSGGAGRLTRGRVQGLLEEEAAARGRLER 298
QY 290 TRIGALALPRGDRPGRAAPAAASARPSKRGGEERVLEKEEEDDEDEDD--VSEG 347
DB 299 TRIGALALPRGDRPGRAAPAAASARAARKRAGEERVLEKEEEDDEDEDDDDVVSEG 358
QY 348 SEVPESDRPAGAQHQLN-GERGQSAKERVKEWTPCGPHQODEGRGAPGSGTRQVFS 406
DB 419 MAALSKEGGSASTTGDSPSPVPLPPGKPALPGADGTPFGCPAGRKEKPADPVEVTMD 478
QY 467 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHHIKVLOOG 526
DB 479 VVEYFTEAGFPEQATAFQOEIDGKSLLMQRTDVLTLGLSIRGLPALKIYEHHIKVLOOG 538
QY 527 HFEDDDPDGFLG 538
DB 539 HFEDDDPEGFLG 550

RESULT 11
US-09-962-055-2
; Sequence 2, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
```

```

; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-962-055-2

Query Match 50.7%; Score 1445.5; DB 9; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 294
Db 11 PKGGAARAGGPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 70

QY 295 LALPRGDRPGRAPPAASARPSRKGGEEVLEKEEEEDDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARAARKRAGEERVLEKEEEEDDEDDDDVVSEGSEVPE 130

QY 353 SDRPAGAHHQLN-GERGPOSAKERVKWTPCPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAHHQLNGGERGPQTAKERAKENSLCGPHQGEGRGPAAGSGTRQVFSMAALS 190

QY 412 KEGGTASVATGPDSPSPVLPFGKPKALPGADGTPFCGCPGRKEKPSDPVMTVMDVVEYF 471
Db 191 KEGGSASSTTGPDSPSPVLPFGKPKALPGADGTPFCGCPAGRKEKPADPVMTVMDVVEYF 250

QY 472 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 531
Db 251 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 310

QY 532 DPGDFLG 538
Db 311 DPEGFLG 317

RESULT 13
US-10-023-529-2
; Sequence 2, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS

```

```

; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-976-740-2

Query Match 50.7%; Score 1445.5; DB 9; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 294
Db 11 PKGGAARAGGPARPVSLREVVRVYLGSGGAGGRLTRGRVQGLLEEEAAARGRLERLTLGA 70

QY 295 LALPRGDRPGRAPPAASARPSRKGGEEVLEKEEEEDDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARAARKRAGEERVLEKEEEEDDEDDDDVVSEGSEVPE 130

QY 353 SDRPAGAHHQLN-GERGPOSAKERVKWTPCPHQGDGEGRPAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAHHQLNGGERGPQTAKERAKENSLCGPHQGEGRGPAAGSGTRQVFSMAALS 190

QY 412 KEGGTASVATGPDSPSPVLPFGKPKALPGADGTPFCGCPGRKEKPSDPVMTVMDVVEYF 471
Db 191 KEGGSASSTTGPDSPSPVLPFGKPKALPGADGTPFCGCPAGRKEKPADPVMTVMDVVEYF 250

QY 472 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 531
Db 251 TEAGPPEQATAFQEQIDGKSLLMORTDVLTLGLSIRLPALKIYEHHIKVLQOQHFEED 310

QY 532 DPGDFLG 538
Db 311 DPEGFLG 317

RESULT 13
US-10-023-529-2
; Sequence 2, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS

```

```
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/10/023,529
/ PRIOR FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: 09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
/ PRIOR FILING DATE: 1996-11-27
/ PRIOR APPLICATION NUMBER: US 60/048,547
/ PRIOR FILING DATE: 1997-06-03
/ NUMBER OF SEQ ID NOS: 53
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 317
/ TYPE: PRT
/ ORGANISM: Oryctolagus cuniculus
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(317)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-023-529-2

Query Match          50.7%; Score 1445.5; DB 13; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVRYLGGSGAGGRLTRGVQGLLEEEAAAARGLERTRLGA 294
Db 11 PKGGAARAGGPARPVSLREVRYLGGSGAGGRLTRGVQGLLEEEAAAARGLERTRLGA 70
QY 295 LALPRGDRPGRAPPAASARPSRKSGSERVLEKEEEDDDDEDEDD--VSEGSVEPE 352
Db 71 LALPRGDRPGRAPPAASARAARKRAGEERVELEKEEEDDDDEDD--VSEGSVEPE 130
QY 353 SDRPAGAQHQLN--GERGQSAKERVKWTGCGPHQGDGEGRPAGSGTROVFSMAANN 411
Db 131 SDRPAGAQHQLNGERGPGQTAKERAKWSLCGPHGQEGRGPAAGSGTROVFSMAALS 190
QY 412 KEGGTASVATGDPSPVPLPGCKPALPGADGTPFCGPKRKEKPSDPVETWMDVVEYF 471
Db 191 KEGGSASSTTGDPSPVPLPGCKPALPGADGTPFCGPKRKEKPADPVETWMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLLMORTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLLMORTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 310
QY 532 DPDGFLG 538
Db 311 DPEGFLG 317

RESULT 14
US-10-023-523-2
/ Sequence 2, Application US/10023523
/ Publication No. US20020152485A1
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Robert S.
/ APPLICANT: Law, Simon W.
/ APPLICANT: Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
/ TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
/ TITLE OF INVENTION: ATHEROSCLEROSIS
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/10/023,523
/ CURRENT FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: US/09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ CURRENT APPLICATION NUMBER: US/10/023,523
/ CURRENT FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: US/09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02

FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/10/023,529
/ PRIOR FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: 09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
/ PRIOR FILING DATE: 1996-11-27
/ PRIOR APPLICATION NUMBER: US 60/048,547
/ PRIOR FILING DATE: 1997-06-03
/ NUMBER OF SEQ ID NOS: 53
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 317
/ TYPE: PRT
/ ORGANISM: Oryctolagus cuniculus
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(317)
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-023-529-2

Query Match          50.7%; Score 1445.5; DB 13; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVRYLGGSGAGGRLTRGVQGLLEEEAAAARGLERTRLGA 294
Db 11 PKGGAARAGGPARPVSLREVRYLGGSGAGGRLTRGVQGLLEEEAAAARGLERTRLGA 70
QY 295 LALPRGDRPGRAPPAASARPSRKSGSERVLEKEEEDDDDEDEDD--VSEGSVEPE 352
Db 71 LALPRGDRPGRAPPAASARAARKRAGEERVELEKEEEDDDDEDD--VSEGSVEPE 130
QY 353 SDRPAGAQHQLN--GERGQSAKERVKWTGCGPHQGDGEGRPAGSGTROVFSMAANN 411
Db 131 SDRPAGAQHQLNGERGPGQTAKERAKWSLCGPHGQEGRGPAAGSGTROVFSMAALS 190
QY 412 KEGGTASVATGDPSPVPLPGCKPALPGADGTPFCGPKRKEKPSDPVETWMDVVEYF 471
Db 191 KEGGSASSTTGDPSPVPLPGCKPALPGADGTPFCGPKRKEKPADPVETWMDVVEYF 250
QY 472 TEAGPPEQATAFQEQEIDGKSLLLMORTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 531
Db 251 TEAGPPEQATAFQEQEIDGKSLLLMORTDVLTLGLSIRLGPALKIYEHHIKVLQOQHFE 310
QY 532 DPDGFLG 538
Db 311 DPEGFLG 317

RESULT 15
US-10-616-187-2
/ Sequence 2, Application US/10616187
/ Publication No. US20040013668A1
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Robert S.
/ APPLICANT: Law, Simon W.
/ APPLICANT: Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
/ TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
/ TITLE OF INVENTION: ATHEROSCLEROSIS
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/10/616,187
/ CURRENT FILING DATE: 2003-07-09
/ PRIOR APPLICATION NUMBER: US/09/616,289
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
/ PRIOR FILING DATE: 1996-11-27
/ PRIOR APPLICATION NUMBER: US 60/048,547
/ PRIOR FILING DATE: 1997-06-03
/ NUMBER OF SEQ ID NOS: 53
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(317)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-616-187-2
```

```
Query Match      50.7%; Score 1445.5; DB 15; Length 317;
Best Local Similarity 89.3%; Pred. No. 5.5e-63;
Matches 274; Conservative 18; Mismatches 12; Indels 3; Gaps 2;

QY 235 PEGGAVRAGGAARPVSLREVYRLGGSGGAGGRLTRGRVQGLLEERAAARGRLERLRLGA 294
Db 11 PKGGAARAGGPARPVSLREVYRLGGSGGAGGRLTRGRVQGLLEERAAARGRLERLRLGA 70

QY 295 LALPRGDRPGRAPPAASARPSRSKRGGEERVLKEKEEEDDEDEDEDD--VSEGSEVPE 352
Db 71 LALPRGDRPGRAPPAASARAARNKRAGEERVLKEKEEEDDEDEDDDDVVSEGSEVPE 130

QY 353 SDRPAGAQHQLN-GERGPOSARVKWTPCGPHQODEGRGAPGSGTRQVFSMAAMN 411
Db 131 SDRPAGAQHQLNGERGPGQTAKERAKENSLCGPHQEEGRGPAAGSGTRQVFSMAALS 190

QY 412 KEGGTASVATGPDSPSPVPLPGKPALPGADGTPFGCPGKPKPSDPVETVMDVVEYF 471
Db 191 KEGGSASSTGPDSPSPVPLPGKPALPGADGTPFGCPGKPKPADPVETVMDVVEYF 250

QY 472 TEAGFPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQQGHFEDD 531
Db 251 TEAGFPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHHIKVLQQGHFEDD 310

QY 532 DPGFLG 538
Db 311 DPEGFLG 317
```

Search completed: September 20, 2005, 12:53:08  
Job time : 297.987 secs

**THIS PAGE BLANK (USPTO)**



Query Match	100.0.0%;	Score 2850;	DB 2;	Length 538;
Best Local Similarity	100.0.0%;	Pred. NO. 6.9e-96;		
Matches 538;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAGPPALPPPETAAATTA	AAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVRR	60
Db	1	MAGPPALPPPETAAATTA	AAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVRR	60
Qy	61	RHGPEPRTAELEKLIQ	QRAVLRSVYKSGISYRNAARVOPPRRGATPPAPPRARGAPA	120
Db	61	RHGPEPRTAELEKLIQ	QRAVLRSVYKSGISYRNAARVOPPRRGATPPAPPRARGAPA	120
Qy	121	AAAAAAPPPTPAPPP	PAPVAAAAPAPAAAAAATAPSPGPAQPGPRAQRAAPLAAP	180
Db	121	AAAAAAPPPTPAPPP	PAPVAAAAPAPAAAAAATAPSPGPAQPGPRAQRAAPLAAP	180
Qy	181	PPAPAAPPAVAPAG	PRRAPPPVAAAREPPLPPPPQPPAPPPQPPQPPPPQPPPEGGAV	240
Db	181	PPAPAAPPAVAPAG	PRRAPPPVAAAREPPLPPPPQPPAPPPQPPQPPPPQPPPEGGAV	240
Qy	241	RAGGAARPVSLREVR	YLGSGGAGGGLTGRVQGLLEEEAAAARGRLERTLGLHALPRG	300
Db	241	RAGGAARPVSLREVR	YLGSGGAGGGLTGRVQGLLEEEAAAARGRLERTLGLHALPRG	300
Qy	301	DRPCRAPPAASAP	SPSRKRGGEERVLKEKEEEDDEDEDEDDVSEGSVPESDRPAGAQ	360
Db	301	DRPCRAPPAASAP	SPSRKRGGEERVLKEKEEEDDEDEDEDDVSEGSVPESDRPAGAQ	360
Qy	361	HHQLNGERGPQSA	KERYKEWTPCCPHQGGQDEGRGAPAGSGTGRQVFSMAANKKEGGTASVA	420

Db	361	HHOLNGERGQSAKERVKEWTPCGPHQGDGEGRGPAGSGTRQVFSMAAMNKEGGTASVA	420
QY	421	TGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKPSDPVETVMDVVFYFTEAGPPEQA	480
Db	421	TGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKPSDPVETVMDVVFYFTEAGPPEQA	480
QY	481	TAFOEQEIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQCGHFEDDDPDGFLG	538
Db	481	TAFOEQEIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQCGHFEDDDPDGFLG	538
RESULT 2			
Q6SP59			
ID	Q6SP59	PRELIMINARY;	PRT; 550 AA.
AC	Q6SP59		
DT	05-JUL-2004	(TReMBLrel. 27, Created)	
DT	05-JUL-2004	(TReMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TReMBLrel. 27, Last annotation update)	
DE	Atherin.		
OS	Oryctolagus cuniculus (Rabbit)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX	NCBI_TaxID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Aorta;		
RA	Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;		
RL	Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY453841; AAR24088.1; -		
DR	InterPro; IPR001660; SAM.		
DR	Pfam; PF00536; SAM 1; 1.		
DR	SMART; SM00454; SAM; 1.		
DR	PROSITE; P550105; SAM DOMAIN; 1.		
SQ	SEQUENCE 550 AA; 57019 MW; AF1CF2B780D879A7 CRC64;		
Query Match 91.0%; Score 2594; DB 2; Length 550;			
Best Local Similarity 91.1%; Pred. No. 1.2e-86;			
Matches 503; Conservative 17; Mismatches 16; Indels 16; Gaps 8;			
QY	1	MAGPPALPPPTAAATTTAAAASSSAASPHYQEWILDTIDSLRSKARPDLEICRMVR	59
Db	1	MAGPPALPPPTAAATTTAAAASSSAASPHYQEWILDTIDSLRSKARPDLEICRMVR	60
QY	60	RRHGPEPETERAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAPRGAP	119
Db	61	RRHGPEPETERAELEKLIQORAVLRVSYKGSISYRNAARVQPPRRGATPPAPRAPRGAP	120
QY	120	AAAAAAPPTTAPPPPPAPV-AAAAAPARAPR--AAAAATAPPSGPAQPGPRAQRAAP	176
Db	121	--AAAAAPPTTAPPPPPAPVAAAAAPARAPRAAAAAAATAPPSGPAQPGPRAQRAAP	178
QY	177	LAAPPAPAAAPVAPAGPRAPP--AFAAREPLPPPPPPAPPOQOQ-----PPPP	229
Db	179	LAAPPAPAAAPAAAPAGPRAPPAAVAARESPLPPPPPPAPPOQOQPPPPPPQ	238
QY	230	QOPPPGCVNAGGARVPSIREVRYLGGSGAGRLTRGVQGLLEEEAAARGLER	289
Db	239	QOPPPGCVNAGGARVPSIREVRYLGGSGAGRLTRGVQGLLEEEAAARGLER	298
QY	290	TRLGALALPRGDRPGRAPPAASARPSKRGGEERVLEKEEEDDDDEDEEDD--VSG	347
Db	299	TRLGALALPRGDRPGRAPPAASARARNKRGGEERVLEKEEEDDDDDDDVVS	358
QY	348	SEVPESDRPAGAQHQLN--GERGPSAKERVKEWTPCGPHQGDGEGRGPAGSGTRQVFS	406
Db	359	SEVPESDRPAGAQHQLNGERGPGQAKERAKESWLCGPHPGQEGRGPAGSGTRQVFS	418
QY	407	MAAMNKEGGTASVATGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKPSDPVETVMD	466
Db	419	MAALSKEGGSASTTGPDSPSPVLPKPKALPGADTTPGCGPPGRKEKPSDPVETVMD	478
QY	467	VVEYFTEAGFPQEQATFQEQEIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQCG	526
Db	479	VVEYFTEAGFPQEQATFQEQEIDGKSLLMQRTDVLTLGSLRGLPALKIYEHHIKVLQCG	538
QY	527	HPEDDDPDGFLG	538
Db	539	HPEDDDPDGFLG	550
RESULT 3			
Q6PIS7			
ID	Q6PIS7	PRELIMINARY;	PRT; 295 AA.
AC	Q6PIS7		
DT	05-JUL-2004	(TReMBLrel. 27, Created)	
DT	05-JUL-2004	(TReMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TReMBLrel. 27, Last annotation update)	
DE	LOC90378 protein (Fragment).		
GN	Name=LOC90378;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung;		
RA	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung;		
RA	Strausberg R.;		
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC030129; AAH30129.1; -		
DR	InterPro; IPR001660; SAM.		
DR	InterPro; IPR010993; SAM_homology.		
DR	Pfam; PF00536; SAM 1; 1.		
DR	SMART; SM00454; SAM; 1.		
DR	PROSITE; P550105; SAM DOMAIN; 1.		
FT	NON TER		
SQ	SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;		
Query Match 54.7%; Score 1560; DB 2; Length 295;			
Best Local Similarity 100.0%; Pred. No. 1.5e-49;			
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	244	GAAPVSLREVRYLGGSGAGRLTRGVQGLLEEEAAARGRLERLRLGALALPRGDRP	303
Db	1	GAAPVSLREVRYLGGSGAGRLTRGVQGLLEEEAAARGRLERLRLGALALPRGDRP	60
QY	304	GRAPPAASARPSKRGGEERVLEKEEEDDDDEEDDVEGSEVPESDRPAGAQHQ	363
Db	61	GRAPPAASARPSKRGGEERVLEKEEEDDDDEEDDVEGSEVPESDRPAGAQHQ	120
QY	364	LNGERGPOSAKERVKEWTPCGPHQGDGEGRGPAGSGTRQVFSMAAMNKEGGTASVATGP	423
Db	121	LNGERGPOSAKERVKEWTPCGPHQGDGEGRGPAGSGTRQVFSMAAMNKEGGTASVATGP	180

```

QY 424 DSPSPVLPKPGKALPGADGTPGCPGPKRKEKSPDVEWTVMDVVEYFTFAGPPEQATAF 483
DB 181 DSPSPVLPKPGKALPGADGTPGCPGPKRKEKSPDVEWTVMDVVEYFTFAGPPEQATAF 240
QY 484 QEOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538
DB 241 QEOEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 295

RESULT 4
Q6P0R3 PRELIMINARY; PRT; 285 AA.
ID Q6P0R3 AC Q6P0R3 PRELIMINARY; PRT; 285 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AAH65477.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;

Query Match 53.1%; Score 1512; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 7,9e-48;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 VRYLGGSGAGRLTRGRVQGLLEEEAAARGHLERLICALALPGDRPGRAPPAASAR 313
DB 1 VRYLGGSGAGRLTRGRVQGLLEEEAAARGHLERLICALALPGDRPGRAPPAASAR 60
QY 314 PSRSKGGSEVLEKEEEDDEDEDDVSGSEVPESDRPAGQHQLNGERGQSA 373
DB 61 PSRSKGGSEVLEKEEEDDEDEDDVSGSEVPESDRPAGQHQLNGERGQSA 120

```

```

QY 374 KERVKEWTPCGPHQODEGRGPAQSGTQVFSMAAMNKEGGTASVATGPDSPSPVLP 433
DB 121 KERVKEWTPCGPHQODEGRGPAQSGTQVFSMAAMNKEGGTASVATGPDSPSPVLP 180
QY 434 GXPALPGADGTPGCPGPKRKEKSPDVEWTVMDVVEYFTFAGPPEQATAFQEOEIDGKSL 493
DB 181 GXPALPGADGTPGCPGPKRKEKSPDVEWTVMDVVEYFTFAGPPEQATAFQEOEIDGKSL 240
QY 494 LLMQRTDVLTLGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538
DB 241 LLMQRTDVLTLGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 285

RESULT 5
Q961M4 PRELIMINARY; PRT; 236 AA.
ID Q961M4 AC Q961M4 PRELIMINARY; PRT; 236 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE LOC90378 protein (Hypothetical protein) (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Director MCG Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007384; AAH07384.2; -.
DR EMBL; BC080588; AAH080588.1; -.
DR HSP; P39769; 1KW4.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 44.5%; Score 1269; DB 2; Length 236;

```

Db	2436	SAKSDGDRIVEPKAGAPDVRDA--KFGGK--VAGAASGRP---LDDERYKAGRADKC 248
QY	289	-----RTR-----LCAALP-----RGDRPGRA 306
Db	2489	EHTVGDQHRTRSPPGTSSESPASASSAASPAREASSSSSPAGTSPDGARGGPPSAA 2548
QY	307	PPAASARPSRSKRGGEERVLEKEEEDDEDEDDVSEGSSEVPESDRPAGAOHQHQLNG 366
Db	2549	GAARAPPGAPGSGGQRR-----GDGPPAADASRGDQAG--HLPAAP 2589
QY	367	ERGP--QSAKERVKEWTPCG-----PHQQDEGRGA-- 396
Db	2590	VAQPLFSAVLRTKARTSGIRDHSRKSMAVLPVRPVSAPRQTGGLPAAQPEGRPPGGL 2649
QY	397	---PGSGTRQVFSMAANKEGGTASVAT-----GPDSP-----SPVP-----LP 432
Db	2650	EPRPQAOQAOPATAPTGPAAAAAATAAVALPAGVPGGVRRAAGRPAGHLRLHIP 2709
QY	433	PGKPA-----LPGADGTPFGCPGKKEKPSDPV 460
Db	2710	PPQPVQLEGVVVPLPASPETP---APAQTPPSPFL 2742
RESULT 7		
Q6ND96		
ID	Q6ND96	PRELIMINARY; PRT; 689 AA.
AC	O6ND96;	
AD	O6ND96;	
DT	05-JUL-2004 (TREMELrel. 27, Created)	
DT	05-JUL-2004 (TREMELrel. 27, Last sequence update)	
DT	05-JUL-2004 (TREMELrel. 27, Last annotation update)	
DE	Possible OmpA family member precursor.	
GN	OrderedLocusNames=RPA0213;	
OS	Rhodopseudomonas palustris.	
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;	
OC	Bradyrhizobiaceae; Rhodopseudomonas.	
OX	NCBI_TaxID=1076;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CGA009 / ATCC BAA-98;	
RX	PubMed=14704707; DOI=10.1038/nbt923;	
RA	Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,	
RA	Land M.D., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,	
RA	Gibson J.L., Hanson T.E., Bobat C., Torres y Torres J.L., Peres C.,	
RA	Harrison F.H., Gibson J., Harwood C.S.;	
RT	"Complete genome sequence of the metabolically versatile	
RT	photosynthetic bacterium Rhodopseudomonas palustris.";	
RL	Nat. Biotechnol. 22:55-61(2004).	
DR	EMBL; BX572593; CAB25657.1; -.	
DR	GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.	
DR	InterPro; IPR006665; OmpA/MotB.	
DR	Pfam; PF00691; OmpA; 1.	
DR	ProDom; PD000930; OmpA/MotB; 1.	
KW	Complete proteome; Signal.	
FT	SIGNAL 1 27 Potential.	
SQ	SEQUENCE 689 AA; 70745 MW; 63178DA3E579B6B CRC64;	
Query Match 12.9%; Score 369; DB 2; Length 689;		
Best Local Similarity 31.6%; Pred. No. 2.6e-06;		
Matches 120; Conservative 23; Mismatches 133; Indels 104; Gaps 15;		
QY	97	ARVQPPREGATPPAPPRGAPAAAAAAPPPTTAPPPPPAPVAAAAPARAPAAAAA 156
Db	52	AXEAFPPGPRPAPAPKAAAPSAPPPPPAAAPHVAPPPPPAPPRAAPP--PPPPAAAPA 109
QY	157	TAPPSGGAQCGPRAQRAAPLAAPPAPAPVAPAGPRAPPPAVAAAREPLPPPP-- 215
Db	110	PAPKAAREPPPPPPSHSAPP---PPPHAAAPP---PPPAKPSAPPTAAAPRAAPPPA 163
QY	216	-----OPRAPPPQQQP-----PPQPQP-----PEGGAVRAGGAARPVSLRVYLGSGGA 264
Db	164	AAPVRPPAPPAGEAPQRGPPPGVAPVNPVNPNA--APDAKEDAAKQ-----PP 212
QY	265	CGRLTRGVQGLLEBEAARGLRTRIGALALRGRDGRPGRA-----PPAASARP 314

```
Db 213 GER--RGPPFG-----AFCTPNATAPGTMTPPGCAPRRGPPPPPPAAAANPPPPAAATP 263
QY 315 SRSRKGGEERVLEKEBEDEDEDDVSEGEVDES-----DRPAGAQHQLNGSRGP 370
Db 264 APS-----AAQAFTSPANPSGPAVAVAPSRGSRGP 296
QY 371 QSAKERVKWTGCGPHQGDGRGAPGSGTGROVFSMAAMNKEGGTASVATGPDSPSPVP 430
Db 297 QPG-----APAGGPPRPQAGFGAGAG-----PAVAPPPGQGPQVPP 333
QY 431 LPPGKALPGADGTPFCPP 450
Db 334 PQGQPPGPAVAPPAPPPP 353

RESULT 8
Q95JC9 PRELIMINARY; PRT; 676 AA.
AC Q95JC9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RA Zhang Q., Szalay A.A., Kyeiyune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY035849; AK61383.1; -.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005249; F: voltage-gated potassium channel activity; IEA.
DR GO: GO:0006813; P: potassium ion transport; IEA.
DR InterPro: IPR005404; KV3.3 channel.
DR PRINTS: PR01582; KV33CHANNEL.
SQ SEQUENCE 676 AA; 62297 MW; 3008BC41EFD81FC9 CRC64;

Query Match 12.8%; Score 364.5; DB 2; Length 676;
Best Local Similarity 31.7%; Pred. No. 3.7e-06;
Matches 124; Conservative 5; Mismatches 163; Indels 99; Gaps 16;

QY 101 PPRGATP--PAPP--RAPGAP-----AAAAAAPTTPAPPPPPAPVAAAAPARA 148
Db 188 PPPPGPPPPGAPPGAPPPGPPPPGPPPPGPPGAPPPGPPPPGPPPPGPPGAPR 247
QY 149 PRAAAAAATAPPSPGPAQGPRAQRAAPLAAPPAAPPAVAVAPPAG--PRRAPPVAVA- 205
Db 248 P--PGPPPLGPPPGAPPAGAPPPGPPPPGPPPPGPPGAPPPGPPPPGPPPPGPP 305
QY 206 --AREPPLPPPPQ--PAPPOQQOPP--PPQPPPEGGAVRAGGAAPVSLREVRYL 258
Db 306 PGARPPGPPPPGPPPGAPPGARPPGPPPPGPPPPGPPGPA--PPGARPPGPPPPGPP 363
QY 259 GSGGAGGRLTRGVQGLLEEEAARGRLERLGLALALPRGDRPGAPPAASARSK 318
Db 364 PGAPPGARPPPPPPP--FGPPPPGAPPGARPPPPGPPPP 400
QY 319 RGGEERVLEKEBEDEDEDDVSEGEVDES DRPAGAQHQLNGERGPGQSAKERVK 378
Db 401 PG-----PPPGPAPPGARPPPPPPPPPPPPADEPQ--GPAFGDKPKKK 440
QY 379 EWTTCGPHQGDGRGAPGSGTGROVFSMAAMNKEGGTASVATGPDSPSPVFLPGKPAL 438
Db 441 PRPPAGPPPPGPPSPGAP--PGAR-----PPPGPPPPGPPPPG--PAP 480
QY 439 PGA-----DGTFFGCPGGRKEKPSDP 459
Db 481 PGARPPPPPPPPGPPPPGAPPGARPPPPGPP 511
```

## RESULT 9

```
Q89X06 PRELIMINARY; PRT; 745 AA.
ID Q89X06
AC Q89X06
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B1r0521 protein.
GN OrderedLocuNames=blr0521;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USD110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohata M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USD110."
RL DNA Res. 9:189-197(2002).
DR EMBL: AP005936; BAC45786.1; -.
DR HSSP: P07176; 1OAP.
DR GO: GO:0008279; C: outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO: GO:0005199; F: structural constituent of cell wall; IEA.
DR InterPro: IPR006665; OmpA/MotB.
DR Pfam: PF00691; OmpA; 1.
DR PRINTS: PR01218; PSTLEXTENSIN.
DR ProDom: PD000930; OmpA/MotB; 1.
DR Complete proteome.
SQ SEQUENCE 745 AA; 74544 MW; 155EDFC74DBCF6D CRC64;

Query Match 12.6%; Score 360; DB 2; Length 745;
Best Local Similarity 30.7%; Pred. No. 5.7e-06;
Matches 124; Conservative 13; Mismatches 137; Indels 130; Gaps 17;

QY 100 QPPR--RGATPPAPPAPRGPAPAAA-----AAAPPTTPAPPPPPAPVAAAAPARA 148
Db 50 QPPKGPAGAAPPAAPARPAAPPAAPPAAPPAAPPAAPPAAPPAAPPAAPPAAPPA 109
QY 149 PRAAAAAATAPPSPGPAQGPRAQRAAPLAAPPAAPPAVAVAPPAGPRA-----PP 201
Db 110 PPPPPPAAPKQSPPPA--AAAPQCHAPTTPPPAPPAAPPAAPPAAPPAAPPAAP 166
QY 202 PAVAAAREPPLPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 249
Db 167 PPPAPRTTPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 226
QY 250 --SLREVRYLGGSGAGGRLTRGVQGLLEEEAARGRLERLGLALALPRGDRPGAP 307
Db 227 APTATPATPATPATPGTGPAGPAGRG-----APPQVRFSGSP 265
QY 308 -----PAASARPSRSKRGGEERVLEKEBEDEDEDEDEDEDEDEDEDEDEDEDE 360
Db 266 AAGSPAPGATPATTTTAPG-----GTATPPSGRP----- 296
QY 361 HHQLNGERGPGQSAKERVKWTGCGPHQGDGRGAPGSGTGROVFSMAAMNKEGGTASVA 420
Db 297 -----GPAST-----PAPGAATP-----APTATPAPGALTPP 324
QY 421 TGPDSPPSPVLPFG--KPALPGADGTPFGCPP--GRKEKPSDP 459
Db 325 PGRPGAGTTPGQGGTTPAGAPAAAGTP-AAFPQAGGLPARPAAP 367

RESULT 10
Q95JD1
```

ID	Q95JDI	PRELIMINARY;	PRT;	566 AA.
AC	Q95JDI			
DT	01-DEC-2001	(TReMBLrel. 19, Created)		
DT	01-DEC-2001	(TReMBLrel. 19, Last sequence update)		
DT	01-MAR-2004	(TReMBLrel. 26, Last annotation update)		
DE	Basic proline-rich protein.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Parotid gland;			
RA	Zhang Q., Szalay A.A., Kyeiyune-Nyombi E., Sands J.F., Oberg K.C.,			
RA	Tieche J.-M., Leonora J.;			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY035847; AAK61381.1; -			
DR	GO; GO:0005199; F:structural constituent of cell wall; IEA.			
DR	InterPro; IPR003882; Pistil extensin.			
DR	PRINTS; PR01218; PSTEXTENSIN.			
SQ	SEQUENCE 566 AA; 53213 MW; E33B3B5E1BDEB81A CRC64;			
Query Match 12.5%; Score 356.5; DB 2; Length 566;				
Best Local Similarity 30.5%; Pred. No. 6.2e-06;				
Matches 120; Conservative 6; Mismatches 158; Indels 109; Gaps 15;				
QY	101	PPRRGATPPAPPAPRGAP---AAAAAAPPPTTAPPPPPAPVAAAPARAPAAAAAT 157		
DB	231	PPGAGPPPPPP---PPGPPPPGAPPGARPPPGPPPPGPPGPPGPPGPPGPPGPPG 281		
QY	158	APSPGPAOPGPAQAALPAPAPAPVAPAG-----PRRAPPVAVA---AR 207		
DB	282	GPDPGPPPPGPAHGAPPPGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 341		
QY	208	EPPLPPPPPP---PAPPQQQPP-PPQPPPPGAVRAGGAARVSLREVYLGSG 262		
DB	342	PPGPPPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 399		
QY	263	GAGRLTRGRVQGLLBEAAARGLRTRIGALALPRGRAPPAASARPSRSKRGG 322		
DB	400	PGARPPPPGPP---PGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 434		
QY	323	ERVLEKEEEDDEDEDDVSEGVSEVSDRPAQHHLNGERGQSAKERVKEWTP 382		
DB	435	-----PPPGAPPPGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 476		
QY	383	CGHQHQDEGRGP-----APGSGTRQVFSMAAMNKEGGTASVATGDFSPVPLPGKP 436		
DB	477	AGPPPPPPPPGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 515		
QY	437	ALPGA-----DGTFFGCPGRKEKPSDP 459		
DB	516	APPGARPPGPPPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 548		
RESULT 11				
Q6X248				
ID	Q6X248	PRELIMINARY;	PRT;	3204 AA.
AC	Q6X248			
DT	05-JUL-2004	(TReMBLrel. 27, Created)		
DT	05-JUL-2004	(TReMBLrel. 27, Last sequence update)		
DT	05-JUL-2004	(TReMBLrel. 27, Last annotation update)		
DE	UL36 very large tegument protein.			
GN	ORFNames=BHV5-20;			
OS	Bovine herpesvirus 5.			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicellovirus.			
OX	NCBI_TaxID=35244;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SV507/99;			
RX	MEDLINE=22850801; PubMed=12970418;			
RX	DOI=10.1128/JVI.77.19.10339-10347.2003;			
RA	Delhon G., Moraes M.P., Lu Z., Afonso C.L., Flores E.F., Weiblen R.,			
RA	Kutish G.F., Rock D.L.;			
RL	"Genome of bovine herpesvirus 5.;"			
RL	J. Virol. 77:10339-10347(2003).			
DR	EMBL; AY261359; AAR86125.1; -			
DR	InterPro; IPR006928; Herpes teg N.			
DR	InterPro; IPR005210; Herpes_UL36			
DR	InterPro; IPR000228; RNA3'_Term_cycl.			
DR	Pfam; PF04843; Herpes teg N; 1.			
DR	Pfam; PF03586; Herpes UL36; 1.			
DR	PROSITE; PS01287; RTC; UNKNOWN 1.			
SQ	SEQUENCE 3204 AA; 327244 MW; 29F912A82EC57090 CRC64;			
Query Match 12.2%; Score 349; DB 2; Length 3204;				
Best Local Similarity 22.7%; Pred. No. 4.2e-05;				
Matches 158; Conservative 42; Mismatches 187; Indels 310; Gaps 26;				
QY	2	AGPPALPPPETAAATTAASAAASAAASHYQEWILDTIDLSRSKARPDLEICRMVRR 61		
DB	2586	APTFVLRPE---PAPTAPASTRGAPLAAAPAEADA-----PGVER-ARLTER 2634		
QY	62	HGPEPERTRAELEKLIQORAVLRVSYKSGISYRNAARVQPPRRGATPPAPPAPGAPAA 121		
DB	2635	TGP-----RKSMPAALPWRRPLPAAPPAAPRTHPPPAALP---PTA 2672		
QY	122	AAAAAPPTTAPPPAPVAAAAAPAPAAAAATAATPSPGPAQPGPRAQAAPLAAPP 181		
DB	2673	AGAAASAPSPAAPDSRTHPPPAVPPPAATAASP-PAPPPPTTS-LSPTGAPP 2730		
QY	182	PAPAAPPAVA-----PPAGPRAPP- 202		
DB	2731	PPSLPPELAASTAGTGNQTVRRPCARKSMPAALPRRRPPPPFLPPSAGADGPPARG 2790		
QY	203	-----AVAREPPLP-----PPQPAPPQQQP 226		
DB	2791	SGVDAGAAEPAPPLPAVSPGPPAVSPPPQRPSPADAGPPGAAPAPAPAPPAPPAP 2850		
QY	227	PPQPQPPPEGAVAGGAARVSLREVYLGSGGAGRLTRGRVQGLLBEAAARG 286		
DB	2851	PAP-PAPPAPPSQPRADGPAP-----GPII-----AASRN- 2880		
QY	287	LERTRIGALALPRGDRPGRAPPAASARPSRSKRGGERVLEKEEEDDEDEDDVSE 346		
DB	2881	-----VPAAPAPPRAPPTVLAIFA----- 2899		
QY	347	GSEVPESDRPA--GAQHHQLNGERGQSAKERVKEWTPC-----GHQGD----- 390		
DB	2900	----PpKPRAPKGRPGAPAKPAAGAKARRHARAVDPTVIGSGFALSGPGADGEDADF 2955		
QY	391	-----EGRGPAQSGTRQVFSMAAMNKEGGTASVATGDFSPSPVP----- 430		
DB	2956	GVSGMYVPPDPATQAPGAPAPP-----SVAAALPAAARPAATQAPAGAPPPSVAAA 3010		
QY	431	-----LPPGKPA-----LPGA-----DGTFFGCPGR 452		
DB	3011	PPAAARLAPPAQAPPTTRIVPAAPRVVGGRRPTARPMARLAAPRPRESFAAAPPR 3070		
QY	453	-----KEKPSDPVETVMVVEYFTEAGFPEQATAFOE 485		
DB	3071	PWASRVDLHTRDLLADAEARIRTLPAPEPPVDP-----GLDTADSSSDWSEPETEASE 3126		
QY	486	QEI-----DGKSILLMQRTDVL-----TGLSI 507		
DB	3127	TEAREGRGPRDRRARVVVANSLLGRQYLRGTGISV 3163		
RESULT 12				
Q7SF15				
ID	Q7SF15	PRELIMINARY;	PRT;	636 AA.
AC	Q7SF15			
DT	01-MAR-2004	(TReMBLrel. 26, Created)		
DT	01-MAR-2004	(TReMBLrel. 26, Last sequence update)		
DT	01-MAR-2004	(TReMBLrel. 26, Last annotation update)		

```

DE Hypothetical protein.
GN Name=NCU07438.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikov C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothke G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamyselis M., Mauceli E., Bielek C., Rudd S., Friesman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catheide D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA00100025; EAA35389.1; -.
DR InterPro; IPR000697; EVH1.
DR InterPro; IPR011036; PH related.
DR InterPro; IPR001960; WH1.
DR InterPro; IPR003124; WH2.
DR Pfam; PF00568; WH1; 1.
DR Pfam; PF02205; WH2; 1.
KW Hypothetical protein.
SQ SEQUENCE 636 AA; 66201 MW; AE102B813CA938B6 CRC64;

Query Match 12.1%; Score 343.5; DB 2; Length 636;
Best Local Similarity 30.8%; Pred. No. 2e-05;
Matches 128; Conservative 28; Mismatches 139; Indels 121; Gaps 22;

QY 1 MAGPALPPETAAATAAAASSAA-----SPHYQEWLDTIDSLR-----SRKA 47
DB 268 MRAPP--PPPPAAAPRSVSEITPSTSRGPGVPPPPPPARRSGKLDTENHQEPAPPRFA 325
QY 48 RP---DLERICRMVRRH-----GPERTRALEKLIQORAVLRVSVKGSISY- 93
DB 326 VPPPIADAGFAHSDPPRHPTSPAGPPPPPPRPKPTPLE-DODPSQRFSPVPPFTGQRSVP 384
QY 94 -----RNAARVQPPRR--GATPPAPPRAPGAPAAAAA-----APPTP--- 131
DB 385 PPPPSRSVPPPPPNRNSAQPLPPKAPGAPLPPASSRPPPPMLPTRSPAPQAPPLP 444
QY 132 ---APPPPPAPVAAAAPARAAPAAAAATAPPSPGPAQPGRAAPLAAPLAPPAAP 188
DB 445 TSNAPPPPPPLP-ATQAP---PPPLPATSPAPPPPPAPPAAPPLPAAHAPPPPPMP 500
QY 189 AVAPPAGPRAPPVAAAREPPLP-----PPQPPAPQOQQPPPPQPPPPGEGAVRA 242
DB 501 MPAPSGG---APPP-----PPPPPPGGMGGVPPPPPPPPPPGMPPPPPALPVPDGSRSA 552
QY 243 GGAARPVSLREVRYLGGSGAGG-----RLTRGVQGLLEEEAAARGRLRLTGALAL 297
DB 553 -----VLGGITAGGIRALKKVDRSQI-----RDRSGA-AV 582
QY 298 PRGDRPGRAP---PAASARPSRKRKGGEERVL-----EKBEEDDDDEDEED 342
DB 583 PGGSDSGPASSGLPPAGAAFGGG--GGWADALALALQKKEKYSKSDDEDDGDWD 636

Query Match 11.9%; Score 340.5; DB 2; Length 779;
Best Local Similarity 25.1%; Pred. No. 3e-05;
Matches 140; Conservative 35; Mismatches 195; Indels 187; Gaps 21;

QY 4 PPALPPETAAATAAAASSAAASPHYQEWLDTIDSLRSRVARPDLERICRMVRRHG 63
DB 156 PPSAPPPPSAPAPSPAPSPMSRPPKH--ESKTONQNSIPSAPAP-----VPGA 203
QY 64 PEPERTRALEKLIQORAVLRVSVKGSISYRNAARVQP-----PRRGATPPAPPR- 114
DB 204 PPAPPTSAP-----SIPSAPAPPAPPPAPPPMAAAPPPMAAAPPA 246
QY 115 PRGAPAAAAAAPPPTTAPPPPPAPVAAAAP-----YLGSGSGAGGR 146
DB 247 PPAPPAPPPMNSAPSAFPAPPPAPPPMSSAPSI PALNSQSDNSPASNPPIPGGLPFLA 306
QY 147 -----RAPAAAA-----AATAPPSP-GPAQPGPRA 171
DB 307 EINARSEKGVDDKTVSHASNASTSTNRAPSPSPAPPPAPPVASAPSI SPSAPAPPI 366
QY 172 QRA-----APLAAP-PPAPRAPP-AVAPPAGPRAPPAPVAAAREPPLP-PPQPPAP 221
DB 367 SSAPSI SPSAPAPPAPPPAPPPAPPPVAAAAPPAPPPAPPPMVS---PPAPPAPPPAPS 424
QY 222 QQQQPPPPPPQPPPEGGAVRA--GGAARPVSV-----LREVR-----YLGSGSGAGGR 267
DB 425 MTSAPAPAPAPALQAASKMSKSSPPVSSGFLPFLAEIOKKRDRDFVVGDSNYSTK 484
QY 268 LTRGRVQGLLEEEAAARGRLRLTGALALPRGDRPGRAPPAASARPSRKRKGGEERVL 327
DB 485 VEQTDIQG-----AAGSAPTTTSAAPPPPPAGGAGMSFMSE 518
QY 328 KEEEDDDDEDEEDDVS-----EGSEVPESDRPAGAHQHLNGERGPQSAKERVKEWP 382
DB 519 ISKLSKLSHHEPAPSAAPPLPPSAPOIPTSTAPP-----VP 555

```







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 60.5826 Seconds  
(without alignments)  
854.447 Million cell updates/sec

Title: US-10-671-242-43

Perfect score: 2850

Sequence: 1 MAGPPALPPPTAAATTA.....HIKVLQGHFDDPDGFLG 538

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	334.5	11.7	1585	2 T31611	hypothetical prote
2	316	11.1	907	2 E96636	hypothetical prote
3	314	11.0	464	2 S22897	extensin - Volvox
4	311	10.9	505	2 S72273	actin-depolymerizi
5	298.5	10.5	1958	2 B40505	hypothetical prote
6	297.5	10.4	416	1 SKXLAG	dermal gland prote
7	297	10.4	1188	2 S49915	extensin-like prote
8	295	10.4	574	2 T43556	Wiskott-Aldrich sy
9	295	10.4	574	2 T38819	wiskott-aldrich sy
10	294	10.3	317	2 A28996	proline-rich prote
11	293.5	10.3	539	2 T28770	hypothetical prote
12	293.5	10.3	1201	2 G86441	unknown protein [i
13	290.5	10.2	1733	1 B45344	probable nuclear a
14	289.5	10.2	332	1 PIHUB6	salivary proline-r
15	286.5	10.1	1621	2 T15264	hypothetical prote
16	282.5	9.9	839	2 F75518	hypothetical prote
17	279.5	9.8	1870	2 S37671	MHC class III hist
18	279.5	9.8	1872	2 S36152	MHC class III hist
19	278	9.8	660	1 Q0B83	BHLFI protein - hu
20	277.5	9.7	2142	2 B35098	MHC class III hist
21	276.5	9.7	416	2 T34279	hypothetical prote
22	276	9.7	2715	2 T13049	eyelid - fruit fly
23	274	9.6	3149	1 Q0B88	BPLFI protein - hu
24	273.5	9.6	439	2 S51939	chitinase (EC 3.2.
25	271	9.5	708	2 D96711	hypothetical prote
26	270.5	9.5	801	2 T29018	hypothetical prote
27	269.5	9.5	240	2 B24264	proline-rich prote
28	268.5	9.4	1151	2 T18535	high molecular mas
29	268.5	9.4	3534	2 T42567	tegument protein 2

30	264	9.3	589	2 T29299	hypothetical prote
31	263	9.2	760	2 T06291	extensin homolog T
32	262.5	9.2	461	2 T10741	extensin-like prot
33	262	9.2	1460	1 EDBEIF	immediate-early pr
34	261.5	9.2	775	1 EDBE11	immediate-early pr
35	261.5	9.2	969	2 T15446	hypothetical prote
36	260.5	9.1	5262	2 T03454	ALR protein - huma
37	259.5	9.1	309	2 S10889	proline-rich prote
38	259.5	9.1	310	1 PIHUSD	salivary proline-r
39	259.5	9.1	1100	2 JC8033	leukocyte formin p
40	259	9.1	666	2 B70803	hypothetical prote
41	258	9.1	436	2 T17908	proline/lysine-ric
42	257	9.0	599	2 T10798	phosphorin-S - Vo
43	257	9.0	707	2 A46302	PTB-associated spl
44	255.5	9.0	1611	2 T38236	hypothetical prote
45	255.5	9.0	1638	2 A42091	transcription acti

#### ALIGNMENTS

##### RESULT 1

T31611

hypothetical protein Y50E8A.g - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T31611

R;Steward, C.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z21047

A;Accession: T31611

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1585 <WIL>

A;Cross-references: EMBL:AL117200; NID:el549770; PIDN:CAB55050.1; CESP:Y50E8A.g

A;Experimental source: clone Y50E8A

C;Genetics:

A;Gene: CESP:Y50E8A.g

A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match

Best Local Similarity 11.7%; Score 334.5; DB 2; Length 1585;

Matches 124; Conservative 17; Mismatches 155; Indels 175; Gaps 14;

QY	89	GSISYRNAARVOPRRGATPPAPPRAPRCAPA-----A 121
Db	907	GGYSGSAAALPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPSSGGYSGSGSAAAGGGSSGGYSG 966
QY	122	AAAAAAPPPTPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP-----AAAAATAPPSPGPAQPGPRAQAAAP 176
Db	967	GSAAPPP 1026
QY	177	LAAPPAPAPAP-PAVAPP-----AGPRAPPPPPVAAAREPPL 211
Db	1027	PP 1080
QY	212	PP 235
Db	1081	PP 1140
QY	236	-----EGGAVRAGGAARPVSLREVLYLGGSGGAGGRLTRGRVQGLLEBEAARGLERT 290
Db	1141	APAPSSGG--YSGGS-----GSAAGGGSSGGYSG----- 1171
QY	291	RIGALALPGDRPPGRAPPAASARPSKFGSERVLEKEEEDDEDEDEDDVSGSEV 350
Db	1172	--GSAAPPP 1205
QY	351	PSDRPAGAHQLNGERCQSAKERVKWTTCGPHQGDQEGRGAPGSGTRQVFSMAAM 410
Db	1206	-----GSAAGGGSSGGYSGSAAAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 1259
QY	411	NKEGG-----TASVATGDFSPVPLPPGKPA---LPGADGTPGCPFG 451

Db 1260 AAGGGSSGGYTCGSAAPPPPPPPPPPPPPAPAPAPSSGGYSGSSG 1310

RESULT 2

E96636

hypothetical protein T7P1.21 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C/Accession: E96636

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Chung, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96636

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-907 <STO>

A:Cross-references: UNIPROT:Q9C946; GB:AE005173; NID:g6751696; PIDN:AAF27679.1; GSPDB:GN

C:Genetics:

A:Gene: T7P1.21

A:Map position: 1

Query Match 11.1%; Score 316; DB 2; Length 907;

Best Local Similarity 29.4%; Pred.No. 1.1e-06;

Matches 112; Conservative 39; Mismatches 130; Indels 100; Gaps 19;

QY 33 EWLTIDSLRSRKARPDLERICRMVRRHGPPTRAELEKLI---QORAVLRVSYKG 89

Db 343 EDVISVVEIKQKK--DEIESIDVKETE-----ESVNLDEESVVLNGEQDTIMKISSLE 395

QY 90 SIS-----YRNAARVOPRRGATTPAP-----PRARGAPAA-AAAAAPP 128

Db 396 STSESKLNHSEKVENSQLFPFPPPPPPPLPSFIKTASLPLSPPTTPIADIALSMPP 455

QY 129 PTPAPPPPPA--PVAAPAAAPARAAAAAT-----APSPGPAQPGRAQRAAPLAAPP 182

Db 456 PPPPPPPPAVMPKHFAPPPPPPLPPAVMPLKHFAPPPPTTFAFKLGSAPPPPPPP 515

QY 183 AP---AAPP-----AVAPPAGRRAPPVAAAREPLPP-----PPOPPAPPOQ 224

Db 516 LPTTIAAPPPPPPPRAAVAPPPPP---PPPGTAAPPPPPPPPGTQAAPPPPPPPMQN 572

QY 225 Q-----PPPPQP-----QPPPEGAVRAGGAARPVSLREVVRVYLG 259

Db 573 RAPSPPPPMGNSGSGPPPPPPMPMLANGATPPPPPPPMAMANGAAGPPP--PPPRMGW 630

QY 260 GSGAGGLTRGVQGLLEEEAAARGLER--TWLGLAL--ALPRGDRPGRAPPA----- 309

Db 631 ANGAAGPPPPPPGAARSLRPKKAAT--KLKRSITOLGNLYILKGVKGRDPNKTGSGSR 688

QY 310 ---ASAPRSKRKGGEERVLE 327

Db 689 KAGAGSAPAGKGGMADALAE 709

RESULT 3

S22697

extensin - Volvox carteri (fragment)

C:Species: Volvox carteri

C>Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004

C/Accession: S22697; S21006

R:Ertl, H.; Hallmann, A.; Wenzl, S.; Sumper, M.

EMBO J. 11, 2055-2062, 1992

A:Title: A novel extensin that may organize extracellular matrix biogenesis in Volvox ca

A:Reference number: S22697; MUID:92289669; PMID:1600938

A:Accession: S22697

A:Molecule type: mRNA

A:Residues: 1-464 <HAL>

A:Cross-references: UNIPROT:Q41645; EMBL:X65165; NID:g21991; PIDN:CAA46283.1; PID:g21992

C:Keywords: glycoprotein

Query Match 11.0%; Score 314; DB 2; Length 464;

Best Local Similarity 31.1%; Pred.No. 7.6e-07;

Matches 82; Conservative 18; Mismatches 86; Indels 76; Gaps 10;

QY 5 PALPPPTAAATTAATAASSSAASHYQEWILDTIDSLRSRKARPDLERICRMVRRHGP 64

Db 220 PSPPPPRVSTSPPPARVSSPPPP-----ATRSPPP-----RRITSP 257

QY 65 EFERTRAELEKLIQORAVLRVSYKGSISYRNAARVOPRRGATTPAPPRARGAPAAAAA 124

Db 258 SPVLT-----ASPLPKTSPPPPPPPPPPPPPVAS 288

QY 125 AAPPPTP---APPPPPAPVAAAAAPARAAAA-----AATAPPSPGPAQPGRAQRAAP 176

Db 289 PPPPPPPRVSPSPPPPPQPVSSPPPPPPPPPPPPSPSPPPPPPPPPPPPPPPRSP 348

QY 177 LAAPP---PAPAAPPAVAPPAGRRAPPFAVAREPPLPPPPPPAPPPOQOQPP----- 227

Db 349 SPPPPPRSSPSP---PPVVSPPPPPPPRASPPPPPPPPPPPPPPPPPPPPPPPPATAAAN 407

QY 228 PPQOPPPPEGAVRAGG---AARP 248

Db 408 PPSAP-----SRSRAGGPPGLGTRP 427

RESULT 4

S72273

actin-depolymerizing protein N-WASP, brain - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 23-Apr-1998 #sequence\_revision 01-May-1998 #text\_change 09-Jul-2004

C/Accession: S72273

R:Waki, H.; Miura, K.; Takenawa, T.

EMBO J. 15, 5326-5335, 1996

A:Title: N-WASP, a novel actin-depolymerizing protein, regulates the cortical cytoskeleton

A:Reference number: S72273; MUID:97050838; PMID:8895577

A:Accession: S72273

A:Molecule type: mRNA

A:Status: nucleic acid sequence not shown

A:Residues: 1-505 <MIK>

A:Cross-references: UNIPROT:Q95107; EMBL:D67066; NID:g1644231; PIDN:BAAL1082.1; PID:d101

A:Experimental source: brain

Query Match 10.9%; Score 311; DB 2; Length 505;

Best Local Similarity 31.4%; Pred.No. 1.1e-06;

Matches 93; Conservative 32; Mismatches 107; Indels 64; Gaps 13;

QY 74 EKLIQORAVLRVSYK-----GSIYRNAARVQ-----PPRRGATPPAPPRARGAPAAA 122

Db 243 EAQLKDRSTKVIYDFIEKTGGVEAVKVELRQAPPPPPPSRGPPPPPP----- 292

QY 123 AAAAPPTTAPPPPPAPVAAAAAPARAAAAATATPPSPGPAQPGRAQRAAP--LAAPP 181

Db 293 ----PPHSGGPPPPFARGGAPP---PPPSRAPTAAPPPPPPSRPGVAPPNNRMYPP 345

QY 182 PAPAAPPAVAPPAGRRAPPFAVAREPPLPPPOPPAPPPOQOQPPPPPPPPPEGAVR 241

Db 346 PLPALPSSA--ESGPPPPPPPLSVSGSVAPPPPPPPPPPPP-----GPPPPGGLPSGDGHV 399

QY 242 AGCAARPVSLRWRYLGGSGGAGGLTRGRVQGLLEEA---AAGR---LERTLGLAL 295

Db 400 PTPAGSKAALLDQIR-----EGQLKK-----VEQNSRPVSCSGRDALLDQIRQGLQ 446

QY 296 ALPRGDRPGRAPPAASARPSRKRGGEERVLEK-----EEEEDEDEDEED 343

Db 447 LKSVTDAPESTPPAPA--PTSGIVGALMEVMQKRSKAIIHSDDEDEDEDEDED 500

## RESULT 5

B40505  
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)  
C;Species: suid herpesvirus 1  
C;Date: 10-Apr-1992 #sequence\_revision 10-Apr-1992 #text\_change 09-Jul-2004  
C;Accession: B40505  
R;Cheung, A.K.  
J. Virol. 65, 5260-5271, 1991  
A;Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies virus.  
A;Reference number: A40505; MUID:91374576; PMID:1654441  
A;Accession: B40505  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1958 <CHE>  
A;Cross-references: UNIPROT:Q69340; GB:M57505; NID:G334066; PIDN:AAA47468.1; PID:G334068  
C;Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 10.5%; Score 298.5; DB 2; Length 1958;  
Best Local Similarity 25.1%; Pred. No. 1e-05;  
Matches 124; Conservative 31; Mismatches 139; Indels 201; Gaps 24;  
QY 61 RHGPEPRTAELEKLIQORAVLRVSYKGSISYRNAARVOPRRGATPPAPRAPRGAPA 120  
DB 262 RQGPALDAVQTE---LVHQRQLGCR-DVQADGAVHDGPPSSN-SHPGPP-GPSGPST 315  
QY 121 AAAAAPPPTPA-----PPPPAPVAAA-----APARA-----PRAAAA 154  
DB 316 SHODPPGPTPTSTSHHHHHQGPPTSPPSTSSHODPPGGPPSAETHHHHHQDPPGGP 375  
QY 155 AATA-----PSPGPAQGPGR-----AQRAPLAAPPAAPPAVAPGAPR----- 197  
DB 376 PSTSSHHHHQDPPGGGPPSPPPRSTSSSSHQGPPSTRPPPPQRPWRPPSPQKISE 435  
QY 198 -----GKR-----RAPPVAAAREPLP 212  
DB 436 TRAGSENTAQLFSHSENKLFSPHMGEGGEGDRTGAGGDRDDPPSP-----PPRP 489  
QY 213 PPQPPAPPOQPPPPPPQPPPPGGAVRAGGAARPVSLREVYLYGGSGGAGRLTRGR 272  
DB 490 PPLPPPPP-----PPPPPPQPPPPAGGSARR-----RRRGGGGPPGR--GGR 529  
QY 273 VQGLLEEAARGLETRLGALALPRGDRPGRAPPAASARPSPSRKGGSERVLEKEEEE 332  
DB 530 RRG-----GKR-----RRAEGTEAAAADAEEEEE 552  
QY 333 DDEDEDEDDVSEGS-VPSDRPAG-----AQHQLNGRGPQSAKER----- 376  
DB 553 DGEDEDEDAEDGEGGEGRAGGGAGESESESRAGAPKSAEQQGVGAVLG 612  
QY 377 ---VKEWTPCGPH-QGQDEGRGAPGSGTRQVFSMAAMNKE---GCTASVATGPDSPSP 428  
DB 613 LLVVRD---GLHLDGPERAAGPA-----VAAAEADLLHRCRVLPVLGP----- 653  
QY 429 VPLPPGKPALPGADG 443  
DB 654 -PGARGPVLHGAAG 667

## RESULT 6

SKL1AG  
dermal gland protein APEG precursor - African clawed frog (fragment)  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: S07498; A34140  
R;Gmachl, M.; Berger, H.; Thalhammer, J.; Kreil, G.  
FEBS Lett. 260, 145-148, 1990  
A;Title: Dermal glands of Xenopus laevis contain a polypeptide with a highly repetitive  
A;Reference number: A34140; MUID:90127399; PMID:2298293  
A;Accession: S07498  
A;Molecule type: mRNA  
A;Residues: 1-416 <GNA>  
A;Cross-references: UNIPROT:P17437; EMBL:X51394; NID:G64547; PIDN:CAA35759.1; PID:G13346  
A;Note: the authors translated the codon TTC for residue 8 as Pro

C;Superfamily: dermal gland protein APEG; trefoil homology  
C;Keywords: skin; tandem repeat  
F;1-20/Domain: signal sequence (fragment) #status predicted <SIG>  
F;21-416/Product: dermal gland protein APEG #status predicted <MAT>  
F;23-344/Region: alanine/glutamic acid/glycine/proline-rich repeats  
F;349-389/Domain: trefoil homology <TRF>  
F;349-375,359-374,369-386/Disulfide bonds: #status predicted

Query Match 10.4%; Score 297.5; DB 1; Length 416;  
Best Local Similarity 28.8%; Pred. No. 3.4e-06;  
Matches 121; Conservative 25; Mismatches 181; Indels 93; Gaps 16;

QY 82 VLRSYKGSISYRNAARVOPRRGATP-----PAPPRAPRGAPAAAAAAPTTPAPP 134  
DB 12 ILSVMY---IQQDAGGEPAPAEVAPAPAEAGGAPAPAEAGGAPAPAEAGGAPAPAEAG 68  
QY 135 PPAPVAAAAPAPAPRAAAAAATAP-----PSGPAQPGPRAAAPLAAPPAPAPAPPA 189  
DB 69 AEPAPADGGAPAPAEAGGAPAPAEAGGAPAPAEAGGAPAPAEAGGAPAPAEAGGAPAE 126  
QY 190 VAPPAGPRRAPPAVAAREPPLPPPPPPAP-PQOQPPPP-----QPQPPPEGGAVRA 242  
DB 127 PAPAPAEAGAPAPAEAGAPAEAGAPAEVAPAEVAPAPAEAGAPAPAEAGAPAE 186  
QY 243 GGAARPVSLREVYLYGGSGGAGRLTRGRVQGLLEEAARGRLERLGLALALPRGDR 302  
DB 187 PAPA-----EGEAPAPAPAE-----GEA 204  
QY 303 PGRAP-PAASARPSRSKRGGEERVLKEEEDDEDEDDVSESEVDESRRPACAQH 361  
DB 205 PAPAPAEAGAPAPAEAGAPAPAEAGAPAPAEAGAPAPAEAGAPAPAEAGAPAPAE- 263  
QY 362 HOLNGER-GPOSAKERVKWTPC---GPHQDQEGRGPAPGSGTRQVFSMAAMNKEGTA 417  
DB 264 ----GAAPAPAEAGAPAPAEAGAPAPAEAGAPAPAEAGAPAPAEAGAPAPAE 315  
QY 418 SVATGPDSPSPVLPKPKPALPGADGTPFGCPGRKEKPSDPSDPSVMTVDVYFTEAGFP 477  
DB 316 SPAEG-GAPAAAAPAEAGAPAPAPAEAGAPAPAEAGAPAPAEAGAPAPAEAGAPAPAE 362

## RESULT 7

S49915  
extensin-like protein - maize  
C;Species: Zea mays (maize)  
C;Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C;Accession: S49915  
R;Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A.  
submitted to the EMBL Data Library, June 1994  
A;Description: Pex genes: pollen-specific genes with extensin-like domains.  
A;Reference number: S49915  
A;Accession: S49915  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1188 <RUB>  
A;Cross-references: UNIPROT:Q41805; EMBL:Z34465; NID:G600117; PIDN:CAA84230.1; PID:G60011

Query Match 10.4%; Score 297; DB 2; Length 1188;  
Best Local Similarity 26.2%; Pred. No. 8.2e-06;  
Matches 136; Conservative 36; Mismatches 171; Indels 176; Gaps 27;  
QY 3 GPPALPPPTAAATTTAAASSSAASPHYQEWILDTIDSLRSRKARPDLERICRMVRRRH 62  
DB 415 GPSSSVPV-----CKPAASAPMPSPH-----TPPD-----VSPEP 444  
QY 63 GPEPETRAELEKLIQO-----RAVLRSVYKGSISYRNAARVQPPRRGATPP--- 109  
DB 445 LPEPSVPAPAPMPMTPHSPDADDVVPPTPPVPGKSPDPSQVQPPQPPAASTPPPSLV 504  
QY 110 --APPAPRGA-----PAAAAAAPP-----TPAP-----PPPPAPVAA--APAR 147  
DB 505 KLSPPQAPVGSPPPPVKTTPPPAPIGSPSPPPPPVSVVSPPPPPVKSPPPPAPVGSPPPEK 564



\_\_\_\_\_

2

[illegible]

Db	719	PPPLSKTPAPPPPLSKTPVPPPPGLGRGTSSGPPP-----LGAKGS-----	761
Qy	273	VQGLLSEEAARGLRERTRIGALALPRG-----DRPGRAPPAASARP-----SRSKRGGBE	323
Db	762	-NAPPPPPPPAGRGR-----ASLGLGRGVSVPATAAPKKTALKPLHWSKVTRAAGS--	812
Qy	324	RVLEKEEEEEDEDEEDVSGSEV--PESDRPAGAHQHLNGERGQPSAK-ERVK	378
Db	813	--LWADTQKQENQPRAPEDISELSFSAVSDTTA-----KKSTGRRGSSISKPEKVQ	864
RESULT 13			
B45344			
Probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)			
C;Species: suid herpesvirus 1			
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004			
C;Accession: B45344			
R;Vicek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.			
Virology 179, 365-377, 1990			
A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented OR			
A;Reference number: A45344; MUID:91021039; PMID:2171211			
A;Accession: B45344			
A;Status: translation not shown			
A;Molecule type: DNA			
A;Residues: 1-1733 <VLC>			
A;Cross-references: UNIPROT:P33485; GB:M34651; NID:G334070; PIDN:AAA47471.1; PID:G334072			
C;Superfamily: pseudorabies virus 1 nuclear antigen			
Query Match 10.2%; Score 290.5; DB 1; Length 1733;			
Best Local Similarity 23.7%; Pred. No. 2.1e-05;			
Matches 112; Conservative 27; Mismatches 121; Indels 213; Gaps 18;			
Qy	108	PPAPPRAGR--APAAAAAAPPPTAPP-----PPAPVAAA-----	143
Db	70	PPHPPRPQDHRFTPEARDHRDPRDHLPTTRTRDHQHRPPTTTTITTKDQHPQDPLL	129
Qy	144	-----APAPRAA-----AAAAA-----	170
Db	130	LPTKTLQEDPHLLRPTRDPPSAKTHHHQDPGPGPSTSSHHHQDPGPGPPPPRP	189
Qy	171	-----AQAAPLAAPPAPPAAPVAPPA-----	194
Db	190	STSSSSHQGPSTRPPPPQRPFRPPSPQKISETRAGSENTAQLFSHSENKLFSPH	249
Qy	195	-----GPRAPPAAVAREPPLPPPPAPPQQOQPPPPQPPPPPEG	238
Db	250	MGRGEGDGRGTAGGEGDRDDPPF---PSPPPRPPLPPPP-----PPPPPPPPFAG	301
Qy	239	AVRAGGAARPVSLREVVRVYLGSSGAGGLTRGRVQGLLEEAARGLRTRLGALALP	298
Db	302	SARR-----RRGGPPGRRRGK-----	323
Qy	299	RGDRPGRAPPAASARPSRKRGEERVLEKEEBEEDDEDEDVSEG---SEVPESDR	355
Db	324	-----RRRAEGTEAAAADAEEDGDGEDDEDRAEGREGDGGR	367
Qy	356	PAGAHQHLNGERG-----POSAKER-----VKWTCGPH-OQDGRGPAPG	398
Db	368	GAGGAGESSESSEAGAEAPSAEQVGVAGVLGLLVVRD-----GLHLDGPERAAGPA--	421
Qy	399	SGTRQVFSMAAMNKE-----GGTASVATGDPSPVPLPGKALPGADG	443
Db	422	-----VAAADADDLHRVPVLVAGAGPGARGPRGVLHGAPGGADAGLEG	466
RESULT 14			
PIHUB6			
salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - human			
N;Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H			
C;Species: Homo sapiens (man)			
C;Date: 04-Dec-1986 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004			
C;Accession: B40750; C40750; A40750; C25372; S02128; S02127; A03293; A90502; A91974; A05			
R;Azen, E.A.; Latreille, P.; Niece, R.L.			

Am, J. Hum. Genet. 53, 264-278, 1993	A;Title: PRB1 gene variants coding for length and null polymorphisms among human salivary
A;Reference number: A40750; MUID:93304421; PMID:8317492	A;Accession: B40750
A;Molecule type: DNA	A;Residues: 35-392 <AZE>
A;Cross-references: UNIPROT:P04280; UNIPROT:P04281; UNIPROT:P02811; UNIPROT:Q16038; GB:S6	A;Experimental source: subject C.J. (large allele)
A;Accession: C40750	A;Molecule type: DNA
A;Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-392 <I>	A;Cross-references: GB:S62929
A;Experimental source: subject M.V.O. (large allele)	A;Accession: A40750
A;Molecule type: DNA	A;Residues: 35-183,245-270,'Q',272-392 <AZ3>
A;Cross-references: GB:S62928	A;Experimental source: subject C.J. (medium allele)
A;Note: authors translated the codon CAA for residue 272 as Arg	R;Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985	A;Title: Differential RNA splicing and post-translational cleavages in the human salivary
A;Reference number: A92492; MUID:85289325; PMID:2993301	A;Accession: C25372
A;Molecule type: mRNA	A;Residues: 1-183,245-392 <MAE>
A;Cross-references: GB:K03204; NID:G190485; PIDN:AAA60185.1; PID:G190486	A;Note: alternatively splice forms lacking portions of the repeat region were also found
R;Lyons, K.M.; Stein, J.H.; Smithies, O.	Genetics 120, 267-278, 1988
A;Title: Length polymorphisms in human proline-rich protein genes generated by intragenic	A;Reference number: S02127; MUID:89121440; PMID:2851479
A;Accession: S02128	A;Status: translation not shown
A;Molecule type: DNA	A;Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>
A;Cross-references: EMBL:X07517	A;Accession: S02127
A;Status: translation not shown	A;Molecule type: DNA
A;Residues: 35-183,245-392 <LY2>	A;Cross-references: EMBL:X07516
R;Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P.	Biochemistry 25, 2387-2392, 1986
A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent structu	A;Reference number: A90502; MUID:86243355; PMID:3521730
A;Accession: A03293	A;Molecule type: protein
A;Residues: 17-38,'Ap',41-51,92-148,'R',150-152 <KA2>	A;Note: among nine basic proline-rich peptides isolated from the saliva, this peptide is
A;Accession: A90502	A;Molecule type: protein
A;Residues: 275-336,'S',338-392 <KAU>	R;Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 94, 1991-1999, 1983	A;Title: Further fractionation of basic proline-rich peptides from human parotid saliva
A;Reference number: A91974; MUID:84161824; PMID:6671974	A;Contents: P-H
A;Accession: A91974	A;Molecule type: protein
A;Residues: 'S',338-392 <SAI>	R;Azen, E.; Lyons, K.M.; McGonigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Vanin,
Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984	A;Reference number: A94005; MUID:84298176; PMID:6089212
A;Accession: A05261	A;Molecule type: DNA
A;Residues: 35-39,'P',41-84,'G',86,'R',87-154,'R',218-246;300-306,'T',308-329,'C',331-384	A;Accession: A05262
A;Molecule type: DNA	A;Residues: 'N',57-59,'A',61-69;334-336,'S',338-339,'R',341-392 <AZ5>
R;Kauffman, D.; Wong, R.; Bennick, A.; Keller, P.	Biochemistry 21, 6558-6562, 1982
A;Title: Basic proline-rich proteins from human parotid saliva: complete covalent structu	



A;Reference number: A90464; MUID:83101329; PMID:6924859  
A;Contents: IB-9

A;Accession: A90464

A;Molecule type: protein

A;Residues: 92-127,'R',129-148,'R',150-152 <KA>

R;Isomura, S.; Saich, E.; Sanada, K.

J. Biochem. 91, 2067-2075, 1982

A;Title: Fractionation and characterization of basic proline-rich peptides of human parvovirus P-9

A;Reference number: A91966; MUID:83007119; PMID:7118863

A;Contents: P-E

A;Accession: A91966

A;Molecule type: protein

A;Residues: 92-127,'R',129-148,'R',150-152 <ISE>

C;Comment: This peptide contains 21-residue repeats, two of which have internal 7-residue

C;Genetics:

A;Gene: GDB:PRB1

A;Cross-references: GDB:119511; OMIM:180989

A;Map position: 12p13.2-12p13.2

A;Note: each of the tandem repeats contains a candidate splice acceptor site, and several

C;Superfamily: proline-rich protein

C;Keywords: alternative splicing; duplication; parotid gland; phosphoprotein; pyroglutamate

F;17-51,92-152/Product: basic proline-rich peptide IB-1 #status experimental <IB1>

F;275-392/Product: basic proline-rich peptide P-E #status experimental <PPE>

F;275-335/Product: basic proline-rich peptide IB-6 #status experimental <IB6>

F;337-392/Product: basic proline-rich peptide P-F #status experimental <PPF>

F;17/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental

F;24/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 10.2%; Score 289.5; DB 1; Length 392;

Best Local Similarity 27.0%; Pred. No. 7.1e-06;

Matches 100; Conservative 28; Mismatches 146; Indels 97; Gaps 17;

QY 108 PPAPPRAPGAPAAAAAAPP-----TPAPPAPPAVAAAP-----ARAPAAAAATAP 159

DB 50 PPPPPKQPG-----PPPGKNNKQPGPPPPKQPGPPPGKDKSRSPSPKQPG 100

QY 160 PPSGPAQP-GPRAQAALAAAPPAAPPAVAPPAGPRRAPPAVAAAREPPPLPPPPPP 218

DB 101 PPOGGNQPGPPPPKGP-QGPPQGNKPGPPPPKGPQGGPPQGGKQSGSPSPKQ 159

QY 219 APP-----QQQQPPPP--QPQPPPEGAVRAGGAARPVSLRVYLGSGGAGRLTRG 271

DB 160 GPPPGQGNQPGPPPPKGPQGGPPQGNKPGQPPPP-----G 197

QY 272 RVQGLLEERAAARGLRTRGLALALPRGDR--PGRAPPAASARPSRKRGGEEVLEKE 329

DB 198 KPQG-----PPPGDKSQSPSPKQPGP--PPQG-----NQ 229

QY 330 EEDDDDEDEDDVSEGEVPSDRPAGAQH---HQLNGERGQSAKERVKEWTPCG-P 385

DB 230 PQGPPPPKGPQGGPPQGNRPGPPPPKGPQGGPPQGGKSRSPQPGK-----PQGP 284

QY 386 HQQDDEGRGAPGSGTRQVFSMAAMNKEGTASVATGDSPPSPVPLPPGKPALPGADGTP 445

DB 285 PQGGNQPGPPPPKGPQ-----GPPPGGNK-----PQGGPPKQPGPPAGGSKSQS 334

QY 446 FGCPPGRKEKP 456

DB 335 ARAPPGKQPG 345

RESULT 15

T15264

hypothetical protein F59E12.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T15264

R;Johnson, D.

submitted to the EMBL Data Library, May 1997

A;Description: The sequence of C. elegans cosmid F59E12.

A;Reference number: Z18318

A;Accession: T15264

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1621 <JOH>

A;Cross-references: UNIPROT:O01900; EMBL:AF003386; NID:g2088833; PID:g2088843; PIDN:AAB5/

A;Experimental source: strain Bristol N2; clone F59E12

C;Genetics:

A;Gene: CESP:F59E12.9

A;Map position: 2

A;Introns: 30/3; 55/1; 200/2; 299/2; 327/2; 369/3; 589/3; 860/1; 986/1; 1278/1; 1547/1

Query Match 10.1%; Score 286.5; DB 2; Length 1621;

Best Local Similarity 21.7%; Pred. No. 2.9e-05;

Matches 128; Conservative 42; Mismatches 193; Indels 227; Gaps 22;

QY 4 PPALPPETAAAT-----TAAASSSAASPHYQEWILD----- 37

DB 1126 PPPVPPPTATSTOSQQQFLEGLGMTNEIVADAIRGMVATIPQELLMLSGSGPHIRD 1185

QY 38 -----TIDLSRSRKARPDLCRMVRRRHGPEPERTRAELEKLIQORAVLRVSYKGI 91

DB 1186 RSQALTASVLGTSSLEADMRKLMENRRIQ--MEKRFQADMDEVERRNIRRGVYRPPP 1243

QY 92 SYRNAARVO-----PPRGATPPAP-----PRAPRGAP-- 119

DB 1244 PFHNGTDMERDMERRDHSMMRPRPNHPLLQMDTAPPQNGSDPPQPSSECPAPPKNSSKK 1303

QY 120 -----AAAAAAPPPTAPPPAPVAAAAAPAPRAAAA 154

DB 1304 RNNRGNRRNTNEHIQANHEDSDSFSTSSSPSPPPPPPPPPSDDLTP----- 1354

QY 155 AATAPPSPGPAQGPRAQRAALAAAPPPAPAP--APPAVAPPAGPRRAPPAVAAAREPPLPP 213

DB 1355 ---VPPPPPPPTMSKAPTGVPLPVPPPPPLFSPSMLLP-----PPPLPSEKKNPL 1405

QY 214 P-----PQPPAPPQOQPP-----PPQPPPEGGAVRA 242

DB 1406 PTANNHGTDTNGSDIDRNVFTPPSQNYSMPPVMGHEGSGWRGGMPLGMGR 1465

QY 243 CGAARPVSLRVYLGSGGAGRLTRGRVQGLLEEEAAAGRLRTRGLALALPRGDR 302

DB 1466 GGGPPPRGNNSPMRGMPMPMRGGGPPMRG-----GPPMRGGPPMFRGG---PPG-- 1511

QY 303 PGRAPPAASARPSRKRGGEEVLEKEEEDDEDEDEDDVSEGEVPSDRPAG-- 358

DB 1512 PGRGNFSPMRGS-SMRGG-----PHQGDDEGR-GPAPGSGTRQVFSMAAMNKE 1542

QY 359 AQHQLNGERGQSAKERVKEWTPCG-----PHQGDDEGR-GPAPGSGTRQVFSMAAMNKE 413

DB 1543 QYYHDSPPQNRGP-----PMGGLPPPHGGMNMRGGPPPPRG----- 1578

QY 414 GGTASVATGDSPPSPVPLPPGKPALPGADGTPFGCPGRKEKPSDPEVT 463

DB 1579 -----GSHCQGGPPPLMGGPP--PRLMGMPPPPPPPNGGPPPPPSVEVT 1618

Search completed: September 20, 2005, 12:43:26

Job time : 63.5826 secs

**THIS PAGE BLANK (USP 10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 129.084 Seconds  
(without alignments)  
650.171 Million cell updates/sec

Title: US-10-671-242-7

Perfect score: 1170

Sequence: 1 EERVLEKEBEDDEDEDE.....HIKVLQGHFEDDDPGFLG 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1170	100.0	217	2 AAW49041	Aaw49041 Human low
2	1170	100.0	217	4 AAB82803	Aab82803 Human low
3	1170	100.0	538	4 AAB82806	Aab82806 Human low
4	1161	99.2	241	3 AAB58957	Aab58957 Breast an
5	1047.5	89.5	232	4 AAB82799	Aab82799 Rabbit lo
6	1047.5	89.5	252	4 AAB82800	Aab82800 Rabbit lo
7	1047.5	89.5	317	4 AAB82798	Aab82798 Rabbit lo
8	1047.5	89.5	550	4 AAB82807	Aab82807 Rabbit lo
9	1041.5	89.0	317	2 AAW49038	Aaw49038 Rabbit lo
10	708	60.5	132	7 ADE59158	Ade59158 Human pro
11	350	29.9	316	4 AAU31793	Aau31793 Novel hum
12	312	26.7	61	3 AAB34365	Aab34365 Human sec
13	285.5	24.4	116	6 ADA55289	Ada55289 Human pro
14	285	24.4	102	4 AAM39592	Aam39592 Human tra
15	285	24.4	102	4 ABB50215	Abb50215 Human tra
16	285	24.4	114	4 AAM41378	Aam41378 Human pol
17	284	24.3	108	5 ABP58930	Abp58930 Human cel
18	224	19.1	256	8 ADP22596	Adp22596 Sea-squir
19	169	14.4	319	4 AAG74087	Aag74087 Human col
20	169	14.4	866	5 ABB97494	Abb97494 Novel hum
21	166.5	14.2	894	5 ABP69567	Abp69567 Human pol
22	166.5	14.2	894	6 ABR39927	Abr39927 Human pro
23	153.5	13.1	172	6 ABU96686	Abu96686 Human nuc
24	153.5	13.1	172	8 ADQ96306	Adq96306 T cell ac
25	153.5	13.1	356	7 ADB65549	Adb65549 Human pro

26	153.5	13.1	356	8 ADQ96308	Adq96308 T cell ac
27	153	13.1	1004	2 AAW37856	Aaw37856 Human pol
28	153	13.1	1004	2 AAW52830	Aaw52830 A tumour
29	151.5	12.9	193	4 ABG21033	Abg21033 Novel hum
30	149.5	12.8	446	4 ABB59294	Abb59294 Drosophil
31	149	12.7	1004	8 ADQ17955	Adq17955 Human sof
32	148	12.6	529	6 ADA54110	Ada54110 Human pro
33	146.5	12.5	62	3 AAB34364	Aab34364 Gene 9 hu
34	146.5	12.5	231	2 AAW49876	Aaw49876 Human pol
35	146.5	12.5	231	2 AAW59426	Aaw59426 Homo sapi
36	146.5	12.5	464	6 ADA54979	Ada54979 Human pro
37	146.5	12.5	464	8 ABM80060	Abm80060 Tumour-as
38	143.5	12.3	623	7 ADB65358	Adb65358 Human pro
39	142.5	12.2	151	4 AAG89234	Aag89234 Human sec
40	142.5	12.2	577	2 AAW37394	Aaw37394 Human sex
41	142.5	12.2	577	2 AAW37398	Aaw37398 Human sex
42	142.5	12.2	577	2 AAY18288	Aay18288 Human scm
43	142.5	12.2	591	2 AAW37395	Aaw37395 Human sex
44	142.5	12.2	591	2 AAW37399	Aaw37399 Human sex
45	142.5	12.2	591	2 AAY18289	Aay18289 Human scm

ALIGNMENTS

RESULT 1  
AAW49041  
ID AAW49041 standard; protein; 217 AA.  
XX AAW49041;  
AC AAW49041;  
XX 09-NOV-1998 (first entry)  
DT Human low density lipoprotein binding protein LBP-2.  
DE Human low density lipoprotein binding protein LBP-2.  
XX Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.  
XX Homo sapiens.  
XX Location/Qualifiers  
PH Key  
FT Peptide  
FT Peptide /note= "Claim 2"  
FT Peptide /note= "Claim 2"  
FT Peptide /note= "Claim 2"  
FT Peptide /note= "Claim 2"  
FT Peptide /note= "Claim 2"  
FT Peptide /note= "Claim 2"  
XX WO9823282-A1.  
XX 04-JUN-1998.  
XX 26-NOV-1997; 97WO-US021857.  
XX 27-NOV-1996; 96US-0031930P.  
XX 03-JUN-1997; 97US-0048547P.  
XX (BOST-) BOSTON HEART FOUND INC.  
XX Lees AM, Lees RS, Law SW, Arjona AA;  
XX WPI; 1998-322455/28.  
XX N-PSDB; AAV32838.  
XX Nucleic acid encoding low density lipoprotein binding proteins and  
XX related vectors - transformed cells, proteins, and modulators of binding,  
XX useful for treatment and diagnosis of atherosclerosis and for identifying  
XX subjects at risk.  
XX Claim 1; Fig 7; 47pp; English.

CC This polypeptide comprises novel human low density lipoprotein (LDL)  
 CC binding protein LBP-2 that is capable of binding both native and methyl  
 CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see  
 CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human  
 CC LBPs (see AAV49037-42) are claimed. An abnormality in an aspect of LBP  
 CC metabolism or structure is diagnostic of a risk for atherosclerosis. The  
 CC invention provides: methods for determining if an animal is at risk for  
 CC atherosclerosis (e.g. for prenatal screening); methods for treating  
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to  
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and  
 CC methods for treating a cell having an abnormality in LBP structure or  
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as  
 CC well as recombinant vectors and host cells used to produce recombinant  
 CC LBP  
 XX  
 XX Sequence 217 AA;

Query Match 100.0%; Score 1170; DB 2; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-93;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EERVLEKEEEDDDDEDEDDVSEGSVEPSDRPAGAHQHLNGERGQSAKERVKEWT 60  
 DB 1 EERVLEKEEEDDDDEDEDDVSEGSVEPSDRPAGAHQHLNGERGQSAKERVKEWT 60  
 QY 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120  
 DB 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120  
 QY 121 DGTFFGCPGRKEKPSDPVEWTVMDVVEYFTAGFPEQATAFQEQIDKSLLLMORTDV 180  
 DB 121 DGTFFGCPGRKEKPSDPVEWTVMDVVEYFTAGFPEQATAFQEQIDKSLLLMORTDV 180  
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217  
 DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 2  
 AAB82803  
 ID AAB82803 standard; protein; 217 AA.  
 AC AAB82803;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Human low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US0006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 PA (BOST-) BOSTON HEART FOUND INC.  
 XX  
 PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX  
 DR WPI: 2001-565505/63.  
 DR N-PSDB; AAH26494.  
 XX  
 PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 XX Claim 13(g); Fig 7B; 143pp; English.

XX  
 CC The present sequence is that of the N-terminal portion of novel human low  
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
 CC deduced from an isolated partial cDNA clone (see AAH26494). A full-length  
 CC sequence is given in AAH82806. Human LBP-2 is an example of claimed LBP  
 CC polypeptides of the invention that are capable of binding to native and  
 CC methylated low density lipoproteins. Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs. Methods  
 CC of determining if an animal is at risk for atherosclerosis, and methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed  
 XX  
 XX Sequence 217 AA;

Query Match 100.0%; Score 1170; DB 4; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-93;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EERVLEKEEEDDDDEDEDDVSEGSVEPSDRPAGAHQHLNGERGQSAKERVKEWT 60  
 DB 1 EERVLEKEEEDDDDEDEDDVSEGSVEPSDRPAGAHQHLNGERGQSAKERVKEWT 60  
 QY 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120  
 DB 61 PCGPHQODEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120  
 QY 121 DGTFFGCPGRKEKPSDPVEWTVMDVVEYFTAGFPEQATAFQEQIDKSLLLMORTDV 180  
 DB 121 DGTFFGCPGRKEKPSDPVEWTVMDVVEYFTAGFPEQATAFQEQIDKSLLLMORTDV 180  
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217  
 DB 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217

RESULT 3  
 AAB82806  
 ID AAB82806 standard; protein; 538 AA.  
 AC AAB82806;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Human low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US0006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 PA (BOST-) BOSTON HEART FOUND INC.  
 XX  
 PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX  
 DR WPI: 2001-565505/63.  
 DR N-PSDB; AAH26499.  
 XX  
 PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 XX Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein  
CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the  
CC coding region of isolated genomic DNA (see AA26499). It differs from the  
CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the  
CC presence of an additional 321 amino acids at the N-terminus (the cDNA  
CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP  
CC polypeptides of the invention that are capable of binding to native and  
CC methylated low density lipoproteins. Also claimed are biologically active  
CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
CC as expression vectors, cells and methods of producing the LBPs.  
CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529  
CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of  
CC determining if an animal is at risk for atherosclerosis, and methods for  
CC evaluating an agent for use in treating atherosclerosis, and methods for  
CC treating a cell having an abnormality in structure or metabolism of LBP  
CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
CC nucleic acid, and vaccine compositions, are also claimed  
XX  
SQ Sequence 538 AA;

Query Match 100.0%; Score 1170; DB 4; Length 538;  
Best Local Similarity 100.0%; Pred. No. 1.4e-92;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGPOSAKERVKEWT 60  
DB 322 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGPOSAKERVKEWT 381  
QY 61 PCGPHQGDGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKPALPGA 120  
DB 382 PCGPHQGDGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKPALPGA 441  
QY 121 DGTFFGCPGKPKSDPVEWTVDVVEYFTEAGFPEQATFOEIDGKSLLMQRTDV 180  
DB 442 DGTFFGCPGKPKSDPVEWTVDVVEYFTEAGFPEQATFOEIDGKSLLMQRTDV 501  
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217  
DB 502 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 538

RESULT 4  
AAB58957  
ID AAB58957 standard; protein; 241 AA.  
AC AAB58957;  
XX  
XX 27-MAR-2001 (first entry)  
XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.  
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;  
KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
KW cardiovascular disorder; wound healing; neurological disease.  
XX  
XX Homo sapiens.  
OS  
XX WO200055173-A1.  
PN  
XX 21-SRP-2000.  
PD  
XX 08-MAR-2000; 2000WO-US005881.  
PF  
XX 12-MAR-1999; 99US-0124270P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX

PI Rosen CA, Ruben SM;  
XX WPI; 2000-611515/58.  
DR N-PSDB; AAF21860.  
DR  
XX New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention, treatment  
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
PT neurological diseases.  
XX  
PS Claim 11; Page 1112; 1299pp; English.  
XX  
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
CC associated with breast and ovarian cancer. Included in the invention are  
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
CC isolation and characterisation of the DNA and protein sequences of the  
CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;  
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;  
CC antiinflammatory; antitumor; vulnary; anticonvulsant; antibacterial;  
CC antifungal; antiparasitic and cardiac activity. The polynucleotide and  
CC protein sequences are used in the diagnosis of cancer, particularly  
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
CC and agonists may also be used in the diagnosis, prevention and treatment  
CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC cardiovascular disorders such as myocardial ischaemia; wound healing;  
CC neurological diseases such as cerebral anoxia and epilepsy; and  
CC infectious diseases  
XX  
SQ Sequence 241 AA;  
Query Match 99.2%; Score 1161; DB 3; Length 241;  
Best Local Similarity 99.1%; Pred. No. 3.1e-92;  
Matches 215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGPOSAKERVKEWT 60  
DB 25 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGPOSAKERVKEWT 84  
QY 61 PCGPHQGDGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKPALPGA 120  
DB 85 PCGPHQGDGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKPALPGA 144  
QY 121 DGTFFGCPGKPKSDPVEWTVDVVEYFTEAGFPEQATFOEIDGKSLLMQRTDV 180  
DB 145 DGTFFGCPGKPKSDPVEWTVDVVEYFTEAGFPEQATFOEIDGKSLLMQRTDV 204  
QY 181 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 217  
DB 205 LTGLSIRLGPALKIYEHKIKVLQOQHFEDEDDPDGFLG 241  
RESULT 5  
AAB82799  
ID AAB82799 standard; protein; 232 AA.  
XX AAB82799;  
XX  
XX 12-NOV-2001 (first entry)  
XX Rabbit low density lipoprotein binding protein 2 (LBP-2).  
DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
KW  
OS Oryctolagus cuniculus.  
XX  
XX WO200164874-A2.  
PN  
XX 07-SEP-2001.  
PD

```

XX PF 28-FEB-2001; 2001WO-US006356.
XX PR 02-MAR-2000; 2000US-00517849.
XX PR 14-JUL-2000; 2000US-00616289.
XX PA (BOST-) BOSTON HEART FOUND INC.
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX DR WPI; 2001-565505/63.
XX DR N-PSDB; AAH26489.
XX PT New isolated low density lipoprotein binding polypeptide for treating,
XX PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX PS Claim 13(c); Fig 3; 143pp; English.
XX CC The present sequence is that of a partial sequence of novel rabbit low
CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
CC deduced from an isolated cDNA clone (see AAH26489). Full-length rabbit
CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed
CC polypeptides of the invention, termed LBPs, that are capable of binding
CC to native and methylated low density lipoproteins. Also claimed are
CC biologically active fragments and analogues of LBPs, polynucleotides
CC encoding LBPs, as well as expression vectors, cells and methods of
CC producing the LBPs. Methods of determining if an animal is at risk for
CC atherosclerosis, methods for evaluating an agent for use in treating
CC atherosclerosis, and methods for treating a cell having an abnormality in
CC structure or metabolism of LBP are also claimed, as are pharmaceutical
CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
XX XX
SQ Sequence 232 AA;
Query Match 89.5%; Score 1047.5; DB 4; Length 232;
Best Local Similarity 88.6%; Pred. No. 2e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;
QY 1 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQHQLN-GERGQSAKERYK 57
DB 13 EERVLEKEEEEEDEDEDEDDVVSEGSEVPESDRPAGAQHQLNGRGPGTAKERAK 72
QY 58 EWTPCGPHQGQDEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPAL 117
DB 73 EWSLCGPHPGQEGRGAPGSGTRQVFSMAALSKEGGSASSTTGPDPSPSPVPLPGKPAL 132
QY 118 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 177
DB 133 PGADGTPFGCPGRKEKPADPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 192
QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 193 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 232
RESULT 6
AAB82800
ID AAB82800 standard; protein; 252 AA.
XX AC AAB82800;
XX AC AAB82800;
XX DT 12-NOV-2001 (first entry)
XX DE Rabbit low density lipoprotein binding protein 3 (LBP-3).
XX DE Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
XX KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX OS Oryctolagus cuniculus.
XX PN WO200164874-A2.
XX FH Key Location/Qualifiers

PD 07-SEP-2001.
XX PF 28-FEB-2001; 2001WO-US006356.
XX PR 02-MAR-2000; 2000US-00517849.
XX PR 14-JUL-2000; 2000US-00616289.
XX PA (BOST-) BOSTON HEART FOUND INC.
XX PI Lees AM, Lees RS, Law SW, Arjona AA;
XX DR WPI; 2001-565505/63.
XX DR N-PSDB; AAH26490.
XX PT New isolated low density lipoprotein binding polypeptide for treating,
XX PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX PS Claim 13(d); Fig 4; 143pp; English.
XX CC The present sequence is that of a partial sequence of novel rabbit low
CC density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is
CC deduced from an isolated cDNA clone (see AAH26490). A full-length
CC sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an
CC example of claimed polypeptides of the invention, termed LBPs, that are
CC capable of binding to native and methylated low density lipoproteins.
CC Also claimed are biologically active fragments and analogues of LBPs,
CC polynucleotides encoding LBPs, as well as expression vectors, cells and
CC methods of producing the LBPs. Methods of determining if an animal is at
CC risk for atherosclerosis, methods for evaluating an agent for use in
CC treating atherosclerosis, and methods for treating a cell having an
CC abnormality in structure or metabolism of LBP are also claimed, as are
CC pharmaceutical compositions comprising an LBP polypeptide or nucleic
XX XX
SQ Sequence 252 AA;
Query Match 89.5%; Score 1047.5; DB 4; Length 252;
Best Local Similarity 88.6%; Pred. No. 2.2e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;
QY 1 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQHQLN-GERGQSAKERYK 57
DB 33 EERVLEKEEEEEDEDEDEDDVVSEGSEVPESDRPAGAQHQLNGRGPGTAKERAK 92
QY 58 EWTPCGPHQGQDEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPGKPAL 117
DB 93 EWSLCGPHPGQEGRGAPGSGTRQVFSMAALSKEGGSASSTTGPDPSPSPVPLPGKPAL 152
QY 118 PGADGTPFGCPGRKEKPSDPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 177
DB 153 PGADGTPFGCPGRKEKPADPVETWMDVVEYFTAGPPEQATAFQEQIDGKSLLLMOR 212
QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB 213 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 252
RESULT 7
AAB82798
ID AAB82798 standard; protein; 317 AA.
XX AC AAB82798;
XX AC AAB82798;
XX DT 12-NOV-2001 (first entry)
XX DE Rabbit low density lipoprotein binding protein 2 (LBP-2).
XX DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
XX KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX OS Oryctolagus cuniculus.
XX PN Key Location/Qualifiers

```

```
FT Misc-difference 10
FT /note= "encoded by TAG"
PN WO200164874-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US006356.
XX
XX 02-MAR-2000; 2000US-00517849.
XX
XX 14-JUL-2000; 2000US-00616289.
XX
XX (BOST-) BOSTON HEART FOUND INC.
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX N-PSDB; AAH26488.
XX
PT New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
XX Claim 13(b); Fig 2b; 143pp; English.
XX
XX The present sequence is that of a partial sequence of novel rabbit low
XX density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is
XX deduced from an isolated cDNA clone (see AAH26488). Full-length rabbit
XX LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed
XX polypeptides of the invention, termed LBPs, that are capable of binding
XX to native and methylated low density lipoproteins. Also claimed are
XX biologically active fragments and analogues of LBPs, polynucleotides
XX encoding LBPs, as well as expression vectors, cells and methods of
XX producing the LBPs. Methods of determining if an animal is at risk for
XX atherosclerosis, methods for evaluating an agent for use in treating
XX atherosclerosis, and methods for treating a cell having an abnormality in
XX structure or metabolism of LBP are also claimed, as are pharmaceutical
XX compositions comprising an LBP polypeptide or nucleic acid, and vaccine
XX compositions
XX
XX Sequence 317 AA;

Query Match      89.5%; Score 1047.5; DB 4; Length 317;
Best Local Similarity 88.6%; Pred. No. 2.9e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEDDEDEDEDD--VSEGSEVPESDRPAGAHQHQLN-GERGPOSASERVK 57
DB 98 EERVLEKEEEDDEDEDEDDDDDDVVSEGSEVPESDRPAGAHQHQLNGGERGPQTAKERAK 157
QY 58 EWTPCPHQOQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPAL 117
DB 158 ENSLCGPHQOQDEGRGPAGSGTRQVFSMAALSKEGGSASSTGPDSPSPVPLPPGKPAL 217
QY 118 PGADGTFPGCPGRKPKSPDVMTVDVVEYFTEAGFPQATAFQEQEIDGKSLLMQOR 177
DB 218 PGADGTFPGCPGRKPKSPDVMTVDVVEYFTEAGFPQATAFQEQEIDGKSLLMQOR 277
QY 178 TDVLTGLSIRLGPALKIYEHKIVLQOQHGFEDDDPDGFLG 217
DB 278 TDVLTGLSIRLGPALKIYEHKIVLQOQHGFEDDDPDGFLG 317

RESULT 8
AAB82807
ID AAB82807 standard; protein; 550 AA.
XX
XX AAB82807;
XX
XX 12-NOV-2001 (first entry)
XX
XX Rabbit low density lipoprotein binding protein 2 (LBP-2).
DE
XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
KW

atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
Oryctolagus cuniculus.
WO200164874-A2.
07-SEP-2001.
28-FEB-2001; 2001WO-US006356.
02-MAR-2000; 2000US-00517849.
14-JUL-2000; 2000US-00616289.
(BOST-) BOSTON HEART FOUND INC.
Lees AM, Lees RS, Law SW, Arjona AA;
WPI; 2001-565505/63.
N-PSDB; AAH26500.
New isolated low density lipoprotein binding polypeptide for treating,
diagnosing and/or identifying therapeutic agents for atherosclerosis.
Claim 13(k); Fig 2a; 143pp; English.
The present sequence is that of a full-length sequence of novel rabbit
low density lipoprotein binding protein 2 (LBP-2). The amino acid
sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit
LBP-2 is an example of claimed polypeptides of the invention, termed
LBPs, that are capable of binding to native and methylated low density
lipoproteins. Also claimed are biologically active fragments and
analogues of LBPs, polynucleotides encoding LBPs, as well as expression
vectors, cells and methods of producing the LBPs. Polypeptides having
amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-
18) of the present sequence are claimed. Methods of determining if an
animal is at risk for atherosclerosis, methods for evaluating an agent
for use in treating atherosclerosis, and methods for treating a cell
having an abnormality in structure or metabolism of LBP are also claimed,
as are pharmaceutical compositions comprising an LBP polypeptide or
nucleic acid, and vaccine compositions
Sequence 550 AA;

Query Match      89.5%; Score 1047.5; DB 4; Length 550;
Best Local Similarity 88.6%; Pred. No. 5.7e-82;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEDDEDEDEDD--VSEGSEVPESDRPAGAHQHQLN-GERGPOSASERVK 57
DB 331 EERVLEKEEEDDEDEDEDDDDDDVVSEGSEVPESDRPAGAHQHQLNGGERGPQTAKERAK 390
QY 58 EWTPCPHQOQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPSPVPLPPGKPAL 117
DB 391 ENSLCGPHQOQDEGRGPAGSGTRQVFSMAALSKEGGSASSTGPDSPSPVPLPPGKPAL 450
QY 118 PGADGTFPGCPGRKPKSPDVMTVDVVEYFTEAGFPQATAFQEQEIDGKSLLMQOR 177
DB 451 PGADGTFPGCPGRKPKSPDVMTVDVVEYFTEAGFPQATAFQEQEIDGKSLLMQOR 510
QY 178 TDVLTGLSIRLGPALKIYEHKIVLQOQHGFEDDDPDGFLG 217
DB 511 TDVLTGLSIRLGPALKIYEHKIVLQOQHGFEDDDPDGFLG 550

RESULT 9
AAW49038
ID AAW49038 standard; protein; 317 AA.
XX
XX AAW49038;
XX
XX 09-NOV-1998 (first entry)
XX
XX Rabbit low density lipoprotein binding protein LBP-2.
```





CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 132 AA;

Query Match 60.5%; Score 708; DB 7; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-53;  
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 MAAMNKEGGTASVATGDSPPVLPFGKALPGADGTPGCGPRKEKPSDPEVMTVD 145

Db 1 MAAMNKEGGTASVATGDSPPVLPFGKALPGADGTPGCGPRKEKPSDPEVMTVD 60

QY 146 VVEYFTAGPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLQOG 205

Db 61 VVEYFTAGPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHKIKVLQOG 120

QY 206 HFEDDDPDGFLG 217

Db 121 HFEDDDPDGFLG 132

#### RESULT 11

AAU31793  
 ID AAU31793 standard; protein; 316 AA.

XX AC AAU31793;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #2284.

XX KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic

XX PS vaccination, testing and therapy.

XX PS Claim 20; Page 509; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy. The proteins  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or

CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention

XX Sequence 316 AA;

Query Match 29.9%; Score 350; DB 4; Length 316;  
 Best Local Similarity 38.7%; Pred. No. 7e-22;  
 Matches 103; Conservative 7; Mismatches 68; Indels 88; Gaps 12;

QY 1 EERVLEKEREEDD-DEDEDEDDVSGSEV-PESDRPAQHQLNGERGQSAKERVKE 58

Db 53 EERVLEKEREEDD-DEDEDEDDVSGSEV-PESDRPAQHQLNGERGQSAKERVKE 98

QY 59 WTPCG----PHQGDGSG-RGPAPGSGTRQVFSMAAMNKEGGT-----A 96

Db 99 -TAGGDLRVPRRSRSGPGTAFPGXRAGASPGQRHPPGCSWPQXTRKGEQVFFSPPA 157

QY 97 SVATGPDSPSPVLPFGKALPGADGTPGCGPRKEKPSDPEVMTVDV-VYFTFTEAG 154

Db 158 SVATGPDSPSPVLPFGKALPGADGTPGCGPRKEKPSDPEVMTVDV-VYFTFTEAG 217

QY 155 FPEQATAFQEQEIDGKSLLMQRTDVLTLGLSIRLGPALKIYEHK----- 198

Db 218 FPEQATAF-----RAGNXWQIFAFAAAHRCASPACPSPRAS 254

QY 199 ---IKVLQGHFEDD-----DPDGFL 216

Db 255 PENLRAPHOGASARPLXGMMDPDGFL 280

#### RESULT 12

AAAB34365

ID AAB34365 standard; protein; 61 AA.

XX AC AAB34365;

XX DT 26-JAN-2001 (first entry)

XX DE Human secreted protein sequence encoded by gene 9 SEQ ID NO:126.

XX KW Human; secreted protein; diagnosis; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active general; vulnary;  
 KW gastrointestinal; nephrotropic; antinfecive; gynaecological;  
 KW and antibacterial; gene therapy; detection; cancer; chromosome marker;  
 KW chromosome identification; neural disorder; immune disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW proliferative disorder; wound healing; infectious disease; preservative;  
 KW food additive.

XX OS Homo sapiens.

XX PN WO200056883-A1.

XX PD 28-SEP-2000.

XX PF 16-MAR-2000; 2000WO-US006822.

XX PR 23-MAR-1999; 99US-0126054P.

XX PR 10-DEC-1999; 99US-0169916P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-587666/55.

XX CC Human secreted proteins and gene sequences encoding them, useful for  
 CC detecting, preventing, and treating disorders such as cancer,  
 CC neurological disorders and immune system disorders.

XX PS Disclosure; Page 396-397; 429pp; English.

XX The polynucleotide sequences given in AAC59566 to AAC59614 encode the  
 CC human secreted proteins given in AAB34299 to AAB34347. AAB34348 to  
 CC AAB34437 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Example of activities include:  
 CC neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular  
 CC active general; vulnary; gastrointestinal; nephrotropic; antiinfective;  
 CC gynaecological; and antibacterial. The polynucleotides can be used for  
 CC the detection of various disorders such as cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The secreted proteins can be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wound healing, and infectious diseases. The proteins can also  
 CC be used as a food additive or preservative to increase or decrease  
 CC storage capabilities. AAC59557 to AAC59565 and AAB34298 represent  
 CC sequences used in the exemplification of the present invention

XX Sequence 61 AA;  
 SQ

Query Match 26.7%; Score 312; DB 3; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-19;  
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 PSDPVEVTVMVVVEYFTEAGFPQATAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKI 194  
 DB 1 PSDPVEVTVMVVVEYFTEAGFPQATAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKI 60  
 QY 195 Y 195  
 DB 61 Y 61

RESULT 13  
 ADA55289  
 ID ADA55289 standard; protein; 116 AA.  
 AC ADA55289;  
 DT 20-NOV-2003 (first entry)  
 XX Human protein, SEQ ID 2857.  
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease.  
 OS Homo sapiens.  
 XX EP1293569-A2.  
 XX 19-MAR-2003.  
 XX 21-MAR-2002; 2002EP-00006586.  
 XX 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX (HELI-) HELIX RES INST.  
 PA (HEAS-) RES ASSOC BIOTECHNOLOGY.  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI; 2003-395539/39.  
 DR N-PSDB; ADA53650.  
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 14; SEQ ID NO 2857; 205pp; English.  
 PS The present invention relates to novel human secretory or membrane  
 XX proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX Sequence 116 AA;  
 SQ

Query Match 24.4%; Score 285.5; DB 6; Length 116;  
 Best Local Similarity 46.9%; Pred. No. 8e-17;  
 Matches 60; Conservative 14; Mismatches 26; Indels 28; Gaps 2;

QY 83 VFSMAAMNKEGTSVATGDPSPVPLPPGPKALPGADGTGPGGPKRKEKPSDVEWT 142  
 DB 15 MLSVDMENKENGSGVGKNSMEN-----GR---PPDPADWA 46  
 QY 143 VMDVVVEYFTEAGFPQATAFQEQEIDGKSLLLMQRTDVLTLGLSIRLGPALKIYEHKVL 202  
 DB 47 VMDVVVYFRTVGFEQASAFQEQEIDGKSLLLMRNDVLTGLQLKLGPKALKIYEHVKPL 106  
 QY 203 QCGHFEDD 210  
 DB 107 QTKHLKNV 114

RESULT 14  
 AAM39592  
 ID AAM39592 standard; protein; 102 AA.  
 XX AAM39592;  
 DT 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 2737.  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX Homo sapiens.  
 OS WO200153312-A1.  
 XX 26-JUL-2001.  
 XX 26-DEC-2000; 2000WO-US034263.  
 XX 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00596042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00691036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AAI58748.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX Example 4; SEQ ID NO 2737; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
 CC encoded polypeptides (AA138642-AA42213) with nootropic,  
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 XX part of the printed specification

SQ Sequence 102 AA;

Query Match 24.4%; Score 285; DB 4; Length 102;

Best Local Similarity 46.9%; Pred. No. 7.5e-17;  
 Matches 60; Conservative 14; Mismatches 26; Indels 28; Gaps 2;

QY 83 VFSMAANKEGGTASVATGPDSPVLPKPKALPGADGTFGCGPKRKEKPSDPVEWT 142

DB 1 MLSVDMENKENGSGVGVNSMEN-----GR---PPDPADWA 32

QY 143 VMDVVEYFTAGPPEQATAFQEQEIDGKSLLMORTDVLTLGLSIRLGPALKIYEHKIVL 202

DB 33 VMDVNVYFRTVGPEEQASAFQEQEIDGKSLLMTRNDVLTGLQLKLGPAKIVYHVKPL 92

QY 203 QQGHFEDD 210

DB 93 QTKHLKNN 100

RESULT 15

ABB50215  
 ID ABB50215 standard; protein; 102 AA.

AC ABB50215;

DT 05-FEB-2002 (first entry)

DE Human transcription factor TRFX-66.

XX Human; transcription factor; TRFX; cell proliferative disease;

KW autoimmune disease; inflammation; neurological disease;

KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;

XX neuroprotective; antiinflammatory; gene therapy.

OS Homo sapiens.

XX WO200172777-A2.

XX 04-OCT-2001.

XX 13-MAR-2001; 2001WO-US008117.

XX 13-MAR-2000; 2000US-0188986P.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Baughn MR, Yue H, Lal P, Lu DAM, Patterson C;

PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;

PI Reddy R;

XX WPI; 2001-570896/64.

DR N-PSDB; ABA83039.

XX

PT Novel transcription factor polypeptides, used to treat diseases  
 PT associated with altered activity and expression of TRFX, and to screen  
 PT for agents capable of modulating its activity.

XX Claim 1; Page 210; 327pp; English.

XX The present sequence is the protein sequence for a human transcription  
 CC factor. The transcription factor and its coding sequence are useful in  
 CC the diagnosis, treatment and prevention of diseases associated with  
 CC altered expression of the transcription factor e.g. cell proliferative,  
 CC autoimmune/inflammatory, neurological and developmental disorders. A  
 CC number of specific disorders/diseases are given in the specification,  
 CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,  
 CC allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic  
 CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,  
 CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,  
 CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative  
 CC colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's  
 CC disease, stroke, and viral, bacterial, fungal and protozoal infections

XX SQ Sequence 102 AA;

Query Match 24.4%; Score 285; DB 4; Length 102;

Best Local Similarity 46.9%; Pred. No. 7.5e-17;  
 Matches 60; Conservative 14; Mismatches 26; Indels 28; Gaps 2;

QY 83 VFSMAANKEGGTASVATGPDSPVLPKPKALPGADGTFGCGPKRKEKPSDPVEWT 142

DB 1 MLSVDMENKENGSGVGVNSMEN-----GR---PPDPADWA 32

QY 143 VMDVVEYFTAGPPEQATAFQEQEIDGKSLLMORTDVLTLGLSIRLGPALKIYEHKIVL 202

DB 33 VMDVNVYFRTVGPEEQASAFQEQEIDGKSLLMTRNDVLTGLQLKLGPAKIVYHVKPL 92

QY 203 QQGHFEDD 210

DB 93 QTKHLKNN 100

Search completed: September 20, 2005, 12:34:40

Job time : 131.084 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1170	100.0	217	3	US-08-979-608A-7	Sequence 7, Appli
2	1170	100.0	217	4	US-09-517-849-7	Sequence 7, Appli
3	1170	100.0	217	4	US-09-616-289-7	Sequence 7, Appli
4	1170	100.0	538	4	US-09-616-289-43	Sequence 43, Appli
5	1047.5	89.5	232	3	US-08-979-608A-3	Sequence 3, Appli
6	1047.5	89.5	232	4	US-09-517-849-3	Sequence 3, Appli
7	1047.5	89.5	232	4	US-09-616-289-3	Sequence 3, Appli
8	1047.5	89.5	252	3	US-08-979-608A-4	Sequence 4, Appli
9	1047.5	89.5	252	4	US-09-517-849-4	Sequence 4, Appli
10	1047.5	89.5	252	4	US-09-616-289-4	Sequence 4, Appli
11	1047.5	89.5	317	3	US-08-979-608A-2	Sequence 2, Appli
12	1047.5	89.5	317	4	US-09-517-849-2	Sequence 2, Appli
13	1047.5	89.5	317	4	US-09-616-289-2	Sequence 2, Appli
14	1047.5	89.5	550	4	US-09-616-289-47	Sequence 47, Appli
15	153	13.1	1004	3	US-08-916-352-2	Sequence 2, Appli
16	153	13.1	1072	4	US-09-949-016-8072	Sequence 8072, Ap
17	149	12.7	1004	4	US-09-949-016-6496	Sequence 6496, Ap
18	146.5	12.5	231	3	US-08-974-380-2	Sequence 2, Appli
19	146.5	12.5	231	4	US-09-546-977A-2	Sequence 2, Appli
20	146.5	12.5	231	4	US-09-654-466-2	Sequence 2, Appli
21	146.5	12.5	340	4	US-09-949-016-10568	Sequence 10568, A
22	142.5	12.2	577	2	US-08-852-153-4	Sequence 4, Appli
23	142.5	12.2	591	2	US-08-852-153-6	Sequence 6, Appli
24	142.5	12.2	620	2	US-08-852-153-2	Sequence 2, Appli
25	141.5	12.1	664	2	US-08-852-153-8	Sequence 8, Appli
26	136	11.6	26	3	US-08-979-608A-20	Sequence 20, Appli
27	136	11.6	26	4	US-09-517-849-20	Sequence 20, Appli



APPLICANT: Lees, Robert S.  
APPLICANT: Law, Simon W.  
APPLICANT: Arjona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
TITLE OF INVENTION: ATHEROSCLEROSIS  
FILE REFERENCE: 10797-004001  
CURRENT APPLICATION NUMBER: US/09/616,289  
CURRENT FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 09/517,849  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: US 08/979,608  
PRIOR FILING DATE: 1997-11-26  
PRIOR APPLICATION NUMBER: US 60/031,930  
PRIOR FILING DATE: 1996-11-27  
PRIOR APPLICATION NUMBER: US 60/048,547  
PRIOR FILING DATE: 1997-06-03  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 43  
LENGTH: 538  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-616-289-43

Query Match 100.0%; Score 1170; DB 4; Length 538;  
Best Local Similarity 100.0%; Pred. No. 9.1e-97;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPOSAKERVKWT 60  
DB 322 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPOSAKERVKWT 381  
QY 61 PCGPHQODRGCPAGCGTRQVFSMAAMNKEGTSVATGPDSPVPLPGKPALGA 120  
DB 382 PCGPHQODRGCPAGCGTRQVFSMAAMNKEGTSVATGPDSPVPLPGKPALGA 441  
QY 121 DGTFFGCPGCKEKPSDPPVETVMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQRTDV 180  
DB 442 DGTFFGCPGCKEKPSDPPVETVMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQRTDV 501  
QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217  
DB 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 5  
US-08-979-608A-3  
Sequence 3, Application US/08/979608A  
Patent No. 6355451  
GENERAL INFORMATION:  
APPLICANT: Lees, Ann M.  
Lees, Robert S.  
Law, Simon W.  
Arjona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,608A

FILING DATE: 26-No. 6355451-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/048,547  
FILING DATE: 03-JUN-1997  
APPLICATION NUMBER: US 60/031,930  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 232 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-08-979-608A-3

Query Match 89.5%; Score 1047.5; DB 3; Length 232;  
Best Local Similarity 88.6%; Pred. No. 3.1e-86;  
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPOSAKERVK 57  
DB 13 BERVLEKEEEDDEDEDDVSEGEVPESDRPAQAQHHQLNGRGPOSAKERVK 72  
QY 58 BWTCPGPHQODRGCPAGCGTRQVFSMAAMNKEGTSVATGPDSPVPLPGKPAL 117  
DB 73 ENSLGGPHQEGRGCPAGCGTRQVFSMAALKEGGSASSITGPDSPVPLPGKPAL 132  
QY 118 PGADGTPGCPGCKEKPSDPPVETVMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQ 177  
DB 133 PGADGTPGCPGCKEKPSDPPVETVMDVVEYFTEAGFPEQATAFQEQIDGKSLLMQ 192  
QY 178 TVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217  
DB 193 TVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 232

RESULT 6  
US-09-517-849-3  
Sequence 3, Application US/09517849  
Patent No. 6605588  
GENERAL INFORMATION:  
APPLICANT: Lees, Ann M.  
Lees, Robert S.  
Law, Simon W.  
Arjona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/517,849  
FILING DATE: 02-Mar-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/979,608

```

; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-517-849-3

Query Match      89.5%; Score 1047.5; DB 4; Length 232;
Best Local Similarity 88.6%; Pred. No. 3.1e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEEEDEDEDD--VSEGEVPESDRPAGAQQHQLN-GERGPOSACKERVK 57
Db 13 EERVLEKEEEEEDEDDDDVSEGEVPESDRPAGAQQHQLNCGRGPGQAKERAK 72
QY 58 EWTPCGPHQODGRGPAPGSGTQVFSMAAMKEGGTASVATGPDSPSPVPLPPGKPAL 117
Db 73 EWSLGGPHGQEGRGPAAGSGTQVFSMAALSKEGGSASTTGPDPSPSPVPLPPGKPAL 132
QY 118 PGADGTPFGCPRKEKPSDPVETWMDVVEYFTAGFPPEQATAFQEQIDGKSLLLMOR 177
Db 133 PGADGTPFGCPAGRKEKPADPVETWMDVVEYFTAGFPPEQATAFQEQIDGKSLLLMOR 192
QY 178 TDVLTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
Db 193 TDVLTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 232

RESULT 8
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match      89.5%; Score 1047.5; DB 3; Length 252;
Best Local Similarity 88.6%; Pred. No. 3.4e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEEEDEDEDD--VSEGEVPESDRPAGAQQHQLN-GERGPOSACKERVK 57
Db 33 EERVLEKEEEEEDEDDDDVSEGEVPESDRPAGAQQHQLNCGRGPGQAKERAK 92

```





```
;
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-979-608A-2

Query Match      89.5%; Score 1047.5; DB 3; Length 317;
Best Local Similarity 88.6%; Pred. No. 4.5e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY      1  EERVLEKEEEEDDEDEDEDD--VSEGSEVPESDRPAGAHQLN-GERGQSAKERVK 57
Db      98  EERVLEKEEEEDDEDEDDDDVVSEGSEVPESDRPAGAHQLNGGGRGQTAKERAK 157
QY      58  EWTGCGPHGQDEGRGAPGSGTRQVFSMAANKGGTASVATGPDSPSPVLPKPKPAL 117
Db      158  EWSLGGPHGQDEGRGAPGSGTRQVFSMAALKSGGSASSTTGPDSPSPVLPKPKPAL 217
QY      118  PGADGTPGCGPGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMQR 177
Db      158  EWSLGGPHGQDEGRGAPGSGTRQVFSMAALKSGGSASSTTGPDSPSPVLPKPKPAL 217
QY      118  PGADGTPGCGPGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMQR 177
Db      218  PGADGTPGCGPGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMQR 277
QY      178  TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
Db      278  TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 317

RESULT 12
US-09-517-849-2
; Sequence 2, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING AND TREATING
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
```

```
;
;
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-517-849-2

Query Match      89.5%; Score 1047.5; DB 4; Length 317;
Best Local Similarity 88.6%; Pred. No. 4.5e-86;
Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY      1  EERVLEKEEEEDDEDEDEDD--VSEGSEVPESDRPAGAHQLN-GERGQSAKERVK 57
Db      98  EERVLEKEEEEDDEDEDDDDVVSEGSEVPESDRPAGAHQLNGGGRGQTAKERAK 157
QY      58  EWTGCGPHGQDEGRGAPGSGTRQVFSMAANKGGTASVATGPDSPSPVLPKPKPAL 117
Db      158  EWSLGGPHGQDEGRGAPGSGTRQVFSMAALKSGGSASSTTGPDSPSPVLPKPKPAL 217
QY      118  PGADGTPGCGPGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMQR 177
Db      218  PGADGTPGCGPGRKEKPSDPVETVMDVVEYFTAGPPEQATAFQEQEIDGKSLLLMQR 277
QY      178  TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
Db      278  TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 317

RESULT 13
US-09-616-289-2
; Sequence 2, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING AND TREATING
; BINDING PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
```

Qy 1 EERVLEKEEEEEDEDEDEDD--VSEGSEVPESDRPAGAQHHQLN-GERGPQSAKRVK 57

Search completed: September 20, 2005, 12:45:28  
Job time : 33.9351 secs

---





```

; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-7

Query Match
Best Local Similarity 100.0%; Score 1170; DB 15; Length 217;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 EERVLEKEEEDDEDEDDVSEGVSEVPSDRPAGAQHHQLNGERGPPQSAKERVKEWT 60
Db 1 EERVLEKEEEDDEDEDDVSEGVSEVPSDRPAGAQHHQLNGERGPPQSAKERVKEWT 60

Qy 61 PCGPHQODRGRRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120
Db 61 PCGPHQODRGRRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120

Qy 121 DGPFPFGCPGRKCKPSPDVSWTVMVVEYFTFAGFPQATAFQGEIDGKSLLLMORTDV 180
Db 121 DGPFPFGCPGRKCKPSPDVSWTVMVVEYFTFAGFPQATAFQGEIDGKSLLLMORTDV 180

Qy 181 LTGLSIRLGPAIKIYEHHIKVLQOGHFEDDDPDGFLG 217
Db 181 LTGLSIRLGPAIKIYEHHIKVLQOGHFEDDDPDGFLG 217

RESULT 6
US-10-671-242-7
; Sequence 7, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match
Best Local Similarity 100.0%; Score 1170; DB 15; Length 217;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 EERVLEKEEEDDEDEDDVSEGVSEVPSDRPAGAQHHQLNGERGPPQSAKERVKEWT 60
Db 1 EERVLEKEEEDDEDEDDVSEGVSEVPSDRPAGAQHHQLNGERGPPQSAKERVKEWT 60

Qy 61 PCGPHQODRGRRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPSPVPLPPGKPALPGA 120

```

```
Db 61 PCGPHQDQGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPGKPALPGA 120
QY 121 DGTFFGCGPKRKEKPSDPVWTVMDVVEYFTEAGFPEQATAFQEQIDGKSLLLMORTDV 180
Db 121 DGTFFGCGPKRKEKPSDPVWTVMDVVEYFTEAGFPEQATAFQEQIDGKSLLLMORTDV 180
QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
Db 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217

RESULT 7
US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-43

Query Match 100.0%; Score 1170; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEEEDEDEDEDDVSEGSVEPESDRPAGAHQHLNGERGPQSAKERVKEWT 60
Db 322 EERVLEKEEEEEDEDEDEDDVSEGSVEPESDRPAGAHQHLNGERGPQSAKERVKEWT 381
QY 61 PCGPHQDQGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPGKPALPGA 120
Db 382 PCGPHQDQGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPGKPALPGA 441
QY 121 DGTFFGCGPKRKEKPSDPVWTVMDVVEYFTEAGFPEQATAFQEQIDGKSLLLMORTDV 180
Db 442 DGTFFGCGPKRKEKPSDPVWTVMDVVEYFTEAGFPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
Db 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 8
US-10-023-529-43
; Sequence 43, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-43

Query Match 100.0%; Score 1170; DB 13; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-80;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEEEDEDEDEDDVSEGSVEPESDRPAGAHQHLNGERGPQSAKERVKEWT 60
Db 322 EERVLEKEEEEEDEDEDEDDVSEGSVEPESDRPAGAHQHLNGERGPQSAKERVKEWT 381
QY 61 PCGPHQDQGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPGKPALPGA 120
Db 382 PCGPHQDQGRGAPGSGTRQVFSMAAMNKEGGTASVATGDPSPVPLPGKPALPGA 441
QY 121 DGTFFGCGPKRKEKPSDPVWTVMDVVEYFTEAGFPEQATAFQEQIDGKSLLLMORTDV 180
Db 442 DGTFFGCGPKRKEKPSDPVWTVMDVVEYFTEAGFPEQATAFQEQIDGKSLLLMORTDV 501
QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
Db 502 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 9
US-10-023-523-43
; Sequence 43, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
```



Db	442	DGTPFCGPRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLMQRTDV	501
Qy	181	LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG	217
Db	502	LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG	538
RESULT 11			
US-10-671-242-43			
; Sequence 43, Application US/10671242			
; Publication No. US20040040049A1			
; GENERAL INFORMATION:			
; APPLICANT: Lees, Ann M.			
; APPLICANT: Lees, Robert S.			
; APPLICANT: Law, Simon W.			
; APPLICANT: Arjona, Anibal A.			
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING			
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING			
; FILE REFERENCE: 10797-004001			
; CURRENT APPLICATION NUMBER: US/10/671,242			
; PRIOR FILING DATE: 2003-09-24			
; PRIOR APPLICATION NUMBER: US/09/616,289			
; PRIOR FILING DATE: 2000-07-14			
; PRIOR APPLICATION NUMBER: US 09/517,849			
; PRIOR FILING DATE: 2000-03-02			
; PRIOR APPLICATION NUMBER: US 08/979,608			
; PRIOR FILING DATE: 1997-11-26			
; PRIOR APPLICATION NUMBER: US 60/031,930			
; PRIOR FILING DATE: 1996-11-27			
; PRIOR APPLICATION NUMBER: US 60/048,547			
; PRIOR FILING DATE: 1997-06-03			
; NUMBER OF SEQ ID NOS: 53			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 43			
; LENGTH: 538			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-671-242-43			
Query Match 100.0%; Score 1170; DB 15; Length 538;			
Best Local Similarity 100.0%; Pred. No. 5.1e-80;			
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGQSAKERVKEWT	60
Db	322	EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGQSAKERVKEWT	381
Qy	61	PCGPHQGDGRGPAGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKALPGA	120
Db	382	PCGPHQGDGRGPAGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKALPGA	441
Qy	121	DGTPFCGPRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLMQRTDV	180
Db	442	DGTPFCGPRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLMQRTDV	501
Qy	181	LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG	217
Db	502	LTGLSIRLGPALKIYEHHIKVLQOQHFDDEDDPDGFLG	538
RESULT 10			
US-10-616-187-43			
; Sequence 43, Application US/10616187			
; Publication No. US20040013668A1			
; GENERAL INFORMATION:			
; APPLICANT: Lees, Ann M.			
; APPLICANT: Lees, Robert S.			
; APPLICANT: Law, Simon W.			
; APPLICANT: Arjona, Anibal A.			
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING			
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING			
; FILE REFERENCE: 10797-004001			
; CURRENT APPLICATION NUMBER: US/10/616,187			
; CURRENT FILING DATE: 2003-07-09			
; PRIOR APPLICATION NUMBER: US/09/616,289			
; PRIOR FILING DATE: 2000-07-14			
; PRIOR APPLICATION NUMBER: US 09/517,849			
; PRIOR FILING DATE: 2000-03-02			
; PRIOR APPLICATION NUMBER: US 08/979,608			
; PRIOR FILING DATE: 1997-11-26			
; PRIOR APPLICATION NUMBER: US 60/031,930			
; PRIOR FILING DATE: 1996-11-27			
; PRIOR APPLICATION NUMBER: US 60/048,547			
; PRIOR FILING DATE: 1997-06-03			
; NUMBER OF SEQ ID NOS: 53			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 43			
; LENGTH: 538			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-616-187-43			
Query Match 100.0%; Score 1170; DB 15; Length 538;			
Best Local Similarity 100.0%; Pred. No. 5.1e-80;			
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGQSAKERVKEWT	60
Db	322	EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHLNGERGQSAKERVKEWT	381
Qy	61	PCGPHQGDGRGPAGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKALPGA	120
Db	382	PCGPHQGDGRGPAGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPPGKALPGA	441
Qy	121	DGTPFCGPRKEKPSDPVETVMDVVVEYFTAGPPEQATAFQEQIDGKSLLMQRTDV	180

; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 846  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 665  
 ; LENGTH: 241  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (9)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (122)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-925-298-665

Query Match 99.2%; Score 1161; DB 9; Length 241;  
 Best Local Similarity 99.1%; Pred. No. 9.7e-80;  
 Matches 215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPSAKERVKWT 60  
 Db EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPSAKERVKWT 84  
 QY 61 PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVLPCKPALPGA 120  
 Db PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVLPCKPALPGA 144  
 QY 121 DGTFFCGPGRKEKPSDPVEMTMDVVEYFTAGFPEQATFQEQEIDGSKLLMQRTDV 180  
 Db DGTFFCGPGRKEKPSDPVEMTMDVVEYFTAGFPEQATFQEQEIDGSKLLMQRTDV 204  
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFEDDDDPDGFLG 217  
 Db LTGLSIRLGPALKIYEHHIKVLQOQHFEDDDDPDGFLG 241

RESULT 13  
 US-10-102-806-665  
 ; Sequence 665, Application US/10102806  
 ; Publication No. US20030054421A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA103PIC1  
 ; CURRENT APPLICATION NUMBER: US/10/102,806  
 ; PRIOR FILING DATE: 2002-03-22  
 ; PRIOR APPLICATION NUMBER: 09/925,298  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 846  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 665  
 ; LENGTH: 241  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (9)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (122)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-10-102-806-665

Query Match 99.2%; Score 1161; DB 14; Length 241;  
 Best Local Similarity 99.1%; Pred. No. 9.7e-80;  
 Matches 215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPSAKERVKWT 60  
 Db EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGRGPSAKERVKWT 84  
 QY 61 PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVLPCKPALPGA 120  
 Db PCGPHQGDGRGPAGSGTQVFSMAAMNKEGGTASVATGPDSPVLPCKPALPGA 144  
 QY 121 DGTFFCGPGRKEKPSDPVEMTMDVVEYFTAGFPEQATFQEQEIDGSKLLMQRTDV 180  
 Db DGTFFCGPGRKEKPSDPVEMTMDVVEYFTAGFPEQATFQEQEIDGSKLLMQRTDV 204  
 QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFEDDDDPDGFLG 217  
 Db LTGLSIRLGPALKIYEHHIKVLQOQHFEDDDDPDGFLG 241

RESULT 14  
 US-09-962-055-3  
 ; Sequence 3, Application US/09962055  
 ; Patent No. US20020052033A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lees, Ann M.  
 ; Lees, Robert S.  
 ; Law, Simon W.  
 ; Arjona, Anibal A.  
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
 ; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
 ; TREATING ATHEROSCLEROSIS  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/962,055  
 ; FILING DATE: 24-Sep-2001  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/979,608  
 ; FILING DATE: 26-NOV-1997  
 ; APPLICATION NUMBER: US 60/031,930  
 ; FILING DATE: 27-NOV-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Myers, Louis  
 ; REGISTRATION NUMBER: 35,965  
 ; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617/542-5070  
 ; TELEFAX: 617/542-8906  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 232 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 ; US-09-962-055-3

Query Match 89.5%; Score 1047.5; DB 9; Length 232;  
 Best Local Similarity 88.6%; Pred. No. 3.5e-71;  
 Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;  
 QY 1 EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLN-GRGPGSAKERVK 57  
 Db EERVLEKEEEDDEDEDDVSEGSEVPESDRPAGAHQHLNGGGRGPGTAKERAK 72



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 110.226 Seconds  
(without alignments)  
1008.117 Million cell updates/sec

Title: US-10-671-242-7

Perfect score: 1170

Sequence: 1 BERVLEKBBEEDDEDEDE.....HIKVLQGGHFDDPDGFLG 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1170	100.0	236	2 Q96IM4	Q96IM4 homo sapien
2	1170	100.0	285	2 Q6P0R3	Q6P0R3 homo sapien
3	1170	100.0	295	2 Q6PT57	Q6PT57 homo sapien
4	1170	100.0	538	2 Q6SPF0	Q6SPF0 homo sapien
5	1047.5	89.5	550	2 Q6SPF9	Q6SPF9 oryctolagus
6	192	16.4	863	2 Q66HN3	Q66HN3 rattus norv
7	192	16.4	863	2 Q9JMD2	Q9JMD2 rattus norv
8	170	14.5	863	2 Q9JMD1	Q9JMD1 mus musculu
9	169	14.4	866	2 Q96C73	Q96C73 homo sapien
10	167	14.3	866	2 Q9UHJ3	Q9UHJ3 homo sapien
11	166.5	14.2	904	2 Q9HCF5	Q9HCF5 homo sapien
12	158	13.5	408	2 Q8CC08	Q8CC08 mus musculu
13	158	13.5	408	2 Q8OVG1	Q8OVG1 mus musculu
14	157.5	13.5	675	2 Q6DJ53	Q6DJ53 xenopus tro
15	154.5	13.2	675	2 Q6PAW6	Q6PAW6 xenopus lae
16	153.5	13.1	356	2 Q8N228	Q8N228 homo sapien
17	153	13.1	610	2 Q6N083	Q6N083 homo sapien
18	153	13.1	957	2 Q6GMQ3	Q6GMQ3 homo sapien
19	153	13.1	1004	1 PHC1 HUMAN	P78364 homo sapien
20	151	12.9	1010	2 Q7TT35	Q7TT35 mus musculu
21	151	12.9	1012	1 PHC1 MOUSE	Q64028 mus musculu
22	150	12.8	310	2 Q6TGW0	Q6TGW0 brachydanio
23	149.5	12.8	446	2 Q76866	Q76866 drosophila
24	149.5	12.8	446	2 Q9W4W7	Q9W4W7 drosophila
25	148	12.6	480	2 Q8TB59	Q8TB59 homo sapien
26	148	12.6	496	2 Q8N195	Q8N195 homo sapien
27	148	12.6	529	2 Q96NU1	Q96NU1 homo sapien
28	147	12.6	310	2 Q7ZYX8	Q7ZYX8 brachydanio
29	147	12.6	405	2 Q8BNM8	Q8BNM8 mus musculu
30	147	12.6	761	2 Q6ZPI2	Q6ZPI2 mus musculu
31	147	12.6	858	2 Q64117	Q64117 mus musculu

RESULT 1

Q96IM4	PRELIMINARY;	PRT;	236 AA.
AC	Q96IM4;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)		
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)		
DE	LOC90378 protein (Hypothetical protein) (Fragment).		
GN	Name=LOC90378;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lymph;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Villalon D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lymph;		
RA	Strausberg R.;		
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lymph;		
RA	Director MGC Project;		
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC007384; AAH07384.2; -		
DR	EMBL; BC080588; AAH0588.1; -		
DR	HSSP; P39769; 1KW4.		
DR	InterPro; IPR001660; SAM.		
DR	InterPro; IPR010993; SAM_homology.		
DR	Pfam; PF00536; SAM 1; 1.		
DR	SMART; SM00454; SAM; 1.		

Q8blb7 mus musculu  
Q96bl4 homo sapien  
Q8n306 homo sapien  
Q8k5d9 mus musculu  
Q6nuj6 homo sapien  
Q88463 mus musculu  
Q9qwh1 mus musculu  
Q8ixk0 homo sapien  
Q06635 bovine herp  
Q8ql14 brachydanio  
Q8ql15 brachydanio  
Q8nal9 homo sapien  
Q7qor1 anopheles g  
Q9ukm5 homo sapien

```
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 1170; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 7.2e-63;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEBEEDDEDEDDVSEGSEVPESDRPAGAHHQHLNGRGPQSAKERVKEWT 60
Db 20 EERVLEKEBEEDDEDEDDVSEGSEVPESDRPAGAHHQHLNGRGPQSAKERVKEWT 79
QY 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120
Db 80 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 139
QY 121 DGTFFGCGPGRKEKPSDPVEMTVMDVVEYFTFAGFPEQATAFQEIDGKSLLLMORTDV 180
Db 140 DGTFFGCGPGRKEKPSDPVEMTVMDVVEYFTFAGFPEQATAFQEIDGKSLLLMORTDV 199
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
Db 200 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 236

RESULT 2
Q6P0R3
ID Q6P0R3 PRELIMINARY; PRT; 285 AA.
AC Q6P0R3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AAH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.

KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF11F901 CRC64;

Query Match 100.0%; Score 1170; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 8.6e-63;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEBEEDDEDEDDVSEGSEVPESDRPAGAHHQHLNGRGPQSAKERVKEWT 60
Db 69 EERVLEKEBEEDDEDEDDVSEGSEVPESDRPAGAHHQHLNGRGPQSAKERVKEWT 128
QY 61 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 120
Db 129 PCGPHQGDGGRGAPGSGTRQVFSMAAMNKEGGTASVATGPDSPVPLPGKPALPGA 188
QY 121 DGTFFGCGPGRKEKPSDPVEMTVMDVVEYFTFAGFPEQATAFQEIDGKSLLLMORTDV 180
Db 189 DGTFFGCGPGRKEKPSDPVEMTVMDVVEYFTFAGFPEQATAFQEIDGKSLLLMORTDV 248
QY 181 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 217
Db 249 LTGLSIRLGPALKIYEHHIKVLQOQHFFEDDDPDGFLG 285

RESULT 3
Q6FIS7
ID Q6FIS7 PRELIMINARY; PRT; 295 AA.
AC Q6FIS7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC90378 protein (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
```

```

FT NON TER 1 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;
  Query Match 100.0%; Score 1170; DB 2; Length 295;
  • Best Local Similarity 100.0%; Pred. No. 8.9e-63;
  Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVKEWT 60
DB EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVKEWT 138

QY 61 PCGPHOGDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPALPGA 120
DB PCGPHOGQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPALPGA 198

QY 121 DGTFFGCPGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 180
DB DGTFFGCPGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 258

QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 295

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
AC Q6SPF0;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

  Query Match 100.0%; Score 1170; DB 2; Length 538;
  Best Local Similarity 100.0%; Pred. No. 1.6e-62;
  Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVKEWT 60
DB EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVKEWT 381

QY 61 PCGPHOGDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPALPGA 120
DB PCGPHOGQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPALPGA 441

QY 121 DGTFFGCPGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 180
DB DGTFFGCPGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 501

QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 5
Q6SPF9 PRELIMINARY; PRT; 550 AA.
ID Q6SPF9;
AC Q6SPF9;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

  Query Match 100.0%; Score 1170; DB 2; Length 538;
  Best Local Similarity 100.0%; Pred. No. 1.6e-62;
  Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVKEWT 60
DB EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVKEWT 381

QY 61 PCGPHOGDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPALPGA 120
DB PCGPHOGQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPALPGA 441

QY 121 DGTFFGCPGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 180
DB DGTFFGCPGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 501

QY 181 LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB LTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 538

RESULT 5
Q6SPF9 PRELIMINARY; PRT; 550 AA.
ID Q6SPF9;
AC Q6SPF9;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

  Query Match 89.5%; Score 1047.5; DB 2; Length 550;
  Best Local Similarity 88.6%; Pred. No. 3.6e-55;
  Matches 195; Conservative 15; Mismatches 7; Indels 3; Gaps 2;

QY 1 EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLN-GERGPOSASERVK 57
DB EERVLEKEEEDDEDEDDVSEGEVPSDRPAGAHQHQLNGRGPQSAKERVK 390

QY 58 EWTGPHOGQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPAL 117
DB EWTGPHOGQDEGRGPAGSGTRQVFSMAAMNKEGTSVATGPDSPVPPLPGKPAL 450

QY 118 PGADGTPFCGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 177
DB PGADGTPFCGGRKEKPSDPVWTVMDVVVEYFTEAGFPEQATAFQEQEIDGKSLLMQRTDV 510

QY 178 TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 217
DB TDVLTGLSIRLGPALKIYEHKIKVLOQGHFEDDDPDGFLG 550

RESULT 6
Q66HN3 PRELIMINARY; PRT; 863 AA.
ID Q66HN3;
AC Q66HN3;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Sfmblt1 protein.
GN Name=Sfmblt1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toehlyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

```





```
QY 125 FGCPPGRKE-----KPSDPVWMTVMDVVEYFTAGPPEQATAFQOEIDGKSLLL 174
Db 767 ENKPPSPKEIRIEVDERLHDSNPLKWSADVVRFFIRSTDCAPLARIFLDOEIDGQALLL 826

QY 175 MORTDVLTLGSLRGLPALKIYEHKIVLQOQHFE 208
Db 827 LTLPTVQECMDLKLGPALKLC-HHIERIKFAFYE 859

RESULT 9
Q96C73 PRELIMINARY; PRT; 866 AA.
AC Q96C73;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Scm-like with four mbt domains 1.
GN Name=SPMBT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014614; AAH14614.1; -.
DR HSSP; P39769; 1KW4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IP0004092; MBT.
DR InterPro; IP001660; SAM_2.
DR Pfam; PF02820; MBT; 4.
DR Pfam; PF07647; SAM_2; 1.
DR SMART; SM00561; MBT; 4.
DR SMART; SM00454; SAM; 1.
SQ SEQUENCE 866 AA; 98141 MW; DC67BF35C413B7 CRC64;

Query Match 14.4%; Score 169; DB 2; Length 866;
Best Local Similarity 26.2%; Pred. No. 0.027;
Matches 53; Conservative 32; Mismatches 75; Indels 42; Gaps 6;

QY 10 EDDDEDEDEDDVSGSEVPESDRPAGAHQHLNGLRGFQSAKRVKWTTCGPHQGD 69
Db 700 EDDDDPDEGDDSLRGSTSEQD-----ELOEE-----SRMSEKSKSSSPTQS 744

QY 187 LGPALKIYEHKIVLQOQHFE 208
Db 842 KLGPALKLC-HHIERIKFAFYE 862

RESULT 11
Q9HCF5 PRELIMINARY; PRT; 904 AA.
AC Q9HCF5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE RUI.
GN Name=RUI;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20125026; PubMed=10661410; DOI=10.1016/S1074-7613(00)80163-6;
RA Morel S., Levy F., Bulet-Schiltz O., Brasseur F., Probst-Kepper M.,
RA Peitrequin A.L., Monsarrat B., Van Velthoven R., Cerottini J.C.,
RA Boon T., Gairin J.E., Van den Eynde B.J.;
RT "Processing of some antigens by the standard proteasome but not by the
RT immunoproteasome results in poor presentation by dendritic cells.";
RL Immunity 12:107-117(2000).
DR EMBL; AF168132; AAF19794.1; -.
DR HSSP; P39769; 1KW4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IP0004092; MBT.
DR InterPro; IP001660; SAM.
DR InterPro; IP0011510; SAM_2.
DR Pfam; PF02820; MBT; 4.
DR Pfam; PF07647; SAM_2; 1.
DR SMART; SM00561; MBT; 4.
DR SMART; SM00454; SAM; 1.
SQ SEQUENCE 866 AA; 98229 MW; C65095C458567FC3 CRC64;

Query Match 14.3%; Score 167; DB 2; Length 866;
Best Local Similarity 26.2%; Pred. No. 0.035;
Matches 53; Conservative 31; Mismatches 76; Indels 42; Gaps 6;

QY 10 EDDDEDEDEDDVSGSEVPESDRPAGAHQHLNGLRGFQSAKRVKWTTCGPHQGD 69
Db 700 EDDDDPDEGDDSLRGSTSEQD-----ELOEE-----SRMSEKSKSSSPTQS 744

QY 70 EGRGAPGSGTQVFSMAAMNKEGGTASV---ATGPDSPSPVPLPGKALPGADGTPFG 126
Db 745 EISTSLPPDRQR-----KELRTFFSDDENKPPSPKEIRIEVAE----- 785

QY 127 CPPGRKEKPSDPVWMTVMDVVEYFTAGPPEQATAFQOEIDGKSLLLMORTDVLTLGSL 186
Db 786 ----RLHDSNPLKWSADVVRFFIRSTDCAPLARIFLDOEIDGQALLLTLPTVQECMDL 841

QY 187 LGPALKIYEHKIVLQOQHFE 208
Db 842 KLGPALKLC-HHIERIKFAFYE 862

RESULT 12
Q9HCF5 PRELIMINARY; PRT; 904 AA.
AC Q9HCF5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE RUI.
GN Name=RUI;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20125026; PubMed=10661410; DOI=10.1016/S1074-7613(00)80163-6;
RA Morel S., Levy F., Bulet-Schiltz O., Brasseur F., Probst-Kepper M.,
RA Peitrequin A.L., Monsarrat B., Van Velthoven R., Cerottini J.C.,
RA Boon T., Gairin J.E., Van den Eynde B.J.;
RT "Processing of some antigens by the standard proteasome but not by the
RT immunoproteasome results in poor presentation by dendritic cells.";
RL Immunity 12:107-117(2000).
DR EMBL; AF168132; AAF19794.1; -.
DR HSSP; P39769; 1KW4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IP0004092; MBT.
DR InterPro; IP001660; SAM.
DR InterPro; IP0011510; SAM_2.
DR Pfam; PF02820; MBT; 4.
DR Pfam; PF07647; SAM_2; 1.
DR SMART; SM00561; MBT; 4.
DR SMART; SM00454; SAM; 1.
SQ SEQUENCE 866 AA; 98229 MW; C65095C458567FC3 CRC64;
```

```
AC O9HCF5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KIAA1617 protein (Fragment).
GN Name=KIAA1617;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046837; BAB13443.1; -.
DR HSSP; Q9UQR0; LO11.
DR Genew; HGNC:20256; SFWB2.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR004092; Mbt.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR011510; SAM_2.
DR Pfam; PF02820; Mbt; 4.
DR Pfam; PF07647; SAM_2; 1.
DR SMART; SM00561; Mbt; 4.
DR SMART; SM00454; SAM; 1.
FT NON TER 1 1
SQ SEQUENCE 904 AA; 101666 MW; DB6AB35E4C58E46 CRC64;

Query Match 14.2%; Score 166.5; DB 2; Length 904;
Best Local Similarity 28.9%; Pred. No. 0.04;
Matches 58; Conservative 29; Mismatches 71; Indels 43; Gaps 8;

QY 8 EEEEDDEDEDEDDYSE--GSEVPSSDRPAGQHQLNGERGQPSAKSRVKEWTFCGPH 65
DB 727 EEESEEDAMDODTASEETGSEL--RDDQTDTSSAEVPSARPRRAVTLR----- 774
QY 66 QGQDEGRGAPGSGTRQVFSSAAMNKEGGTASVATGPDSPSPVLPKGPALPGADGTFP 125
DB 775 SGSEPVRRPPPTTRGRGAPASSAEGE-----KCPPTKP--EGTDT-- 817
QY 126 GCPPGRKEK-----PSDPVETVMVDVRYFTAGFPQEQATAPQEQIDGKSLLLMQRTD 179
DB 818 -----KQEEERLVLESNPLEWTVTDVRFIKLTDCAPLAKIFQEQIDGQALLLTLP 872
QY 180 VLTGLSIRLGPALKIYEHNIK 200
DB 873 VQECMELKLGPAIKLC-HQIE 892

RESULT 12
ID Q8CC08 PRELIMINARY; PRT; 408 AA.
AC Q8CC08;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
DE library, clone:933016D17 product:hypothetical SAM domain (Sterile
DE alpha motif) containing protein, full insert sequence.
GN Name=Scml4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
```

```
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Hashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (KISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK034171; BAC28615.1; -.
DR HSSP; P39769; lKw4.
DR MGD; MGI:2446140; Scml4.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000637; A+T_hook.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00384; AT_hook; 1.
DR SMART; SM00454; SAM; 1.
KW Hypothetical protein.
SQ SEQUENCE 408 AA; 44445 MW; 59DD91591D10D59 CRC64;

Query Match 13.5%; Score 158; DB 2; Length 408;
Best Local Similarity 26.9%; Pred. No. 0.059;
```

```

SQ  SEQUENCE 408 AA; 44475 MW; 196FA9ACE1D40C55 CRC64;

Query Match 13.5%; Score 158; DB 2; Length 408;
Best Local Similarity 26.9%; Pred. No. 0.059;
Matches 57; Conservative 33; Mismatches 86; Indels 36; Gaps 9;

Qy 14 DEDEDEDDVSQGSFVPSD-----RPAGAQHQHQLN-GERGPQGAKEKRVKWTPCGPHQGG 68
    |: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 215 DKAQEREDGRTGSAKVATAECLANAVGNRVAMDFSHRGSVTHSSSLYKRLTCG---DS 271

Qy 69 DEGRGPA-PSGSTRQVPSMAANKEGGTASVATGPDSPSPVPLPPCKPALPGADGTPFGC 127
    ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 272 HLAGGPATTTSGR-----TNPVPSGGSSPGLRLPASPSKRGKTAIGNRC 318

Qy 128 PPG-----RKEKSPDPVETVMDVBYFYFTEAGFPE---QATAFQOEIDGKSILLM 175
    :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 319 AFSPSPEVQDTRPPSSRSTWTVEDVRFVKDAD-PEALGPHVELFRKHEIDGNALLL 377

Qy 176 QRTDVLTGSLRGLPALKIYEHKIKVLOQGHF 207
    : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 378 RSDMIMKYLGLKLPALKLC-VHIDIKQAKF 408

RESULT 14
Q6DU53
ID Q6DJ53 PRELIMINARY; PRT; 675 AA.
AC Q6DJ53;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Scmh1-prov protein.
GN Names=scmh1-prov;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8364;
ID [1]
SEQUENCE FROM N.A.
RP TISSUE=Whole body;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smaluk D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
ID [2]
SEQUENCE FROM N.A.
RP TISSUE=Whole body;
RC Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC075330; AAH75330.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0045449; P:regulation of transcription; IEA.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR004092; Mbt.
DR InterPro: IPR011660; SAM.
DR InterPro: IPR011510; SAM.
DR InterPro: IPR011510; SAM.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 24.4357 Seconds  
(without alignments)  
854.447 Million cell updates/sec

Title: US-10-671-242-7

Perfect score: 1170

Sequence: 1 BERVLEKBEEDDEDEDE.....HIKVLQGGHFDDPDGFLG 217

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	12.9	1012	2 I53172	RAE-28 - mouse
2	136	11.6	1300	2 T03166	probable immediate
3	135	11.5	1589	2 T1306	hypothetical prote
4	130	11.1	930	2 D37271	A-alpha 2 4 protei
5	124.5	10.6	2453	2 S60254	nuclear receptor c
6	123.5	10.6	783	2 A55817	cyclin-dependent k
7	123	10.5	300	2 S19560	proline-rich prote
8	123	10.5	754	2 A56619	female sterile hom
9	122.5	10.5	206	1 P1RT3	acidic proline-ric
10	122.5	10.5	295	2 B48013	proline-rich prote
11	121.5	10.4	768	2 H54024	protein kinase (EC
12	121.5	10.4	777	2 F54024	protein kinase (EC
13	121.5	10.4	777	2 B54024	protein kinase (EC
14	121.5	10.4	779	2 B54024	protein kinase (EC
15	119.5	10.2	367	1 OZZQMY	circumsporozoite p
16	118.5	10.1	1477	2 T13797	tumor suppressor p
17	118	10.1	1280	2 T00365	hypothetical prote
18	117.5	10.0	1282	2 JE0120	glycoprotein A - m
19	116	9.9	301	2 E29149	proline-rich prote
20	115.5	9.9	617	2 S42719	actin-binding prot
21	115	9.8	772	2 I50463	protein kinase - c
22	115	9.8	1840	2 T30250	GPI protein - mous
23	114.5	9.8	280	2 S35103	bone stialoprotein
24	114.5	9.8	1460	1 EDBBIF	immediate-early pr
25	114	9.7	909	2 S32538	cGMP-gated cation
26	114	9.7	1388	2 A53317	collagen alpha 1(X
27	113	9.7	226	2 S41032	hypothetical prote
28	113	9.7	916	2 A27864	neurofilament trip
29	112.5	9.6	512	2 E59437	F02569_2 protein {

30	112.5	9.6	647	2 T43952	hypothetical prote
31	112	9.6	786	2 A35466	progesterone recep
32	112	9.6	892	2 B46203	mating type A alph
33	111.5	9.5	1494	2 T14355	protein-tyrosine-p
34	111.5	9.5	1880	2 T18531	tractin - medicina
35	111	9.5	1252	2 T14272	cortactin-binding
36	110	9.4	1110	2 I51116	NP-180 - sea lampr
37	110	9.4	1213	2 S16356	ovo protein - frui
38	109.5	9.4	730	2 A36226	collagen alpha 1 c
39	109	9.3	902	2 T26775	hypothetical prote
40	108	9.2	170	2 A48013	proline-rich prote
41	107.5	9.2	1315	2 A56101	collagen alpha 1(X
42	107.5	9.2	1774	2 B56101	collagen alpha 1(X
43	107	9.1	907	2 A45560	sporozoite surface
44	107	9.1	2924	2 T18378	variant-specific s
45	106.5	9.1	299	2 A29681	nucleophosmin - Af

ALIGNMENTS

RESULT 1

I53172

RAE-28 - mouse

C;Species: Mus sp. (mouse)

C;Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text\_change 23-Sep-2002

C;Accession: I53172; I66850; I66851

R;Nomura, M.; Takihara, Y.; Shimada, K.

Differentiation 57, 39-50, 1994

A;Title: Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cell

Drosophila polyhomeotic protein.

A;Reference number: I53172; MUID:94350162; PMID:8070621

A;Accession: I53172

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1012 <RES>

A;Cross-references: GB:S73882; NID:G688186; PIDN:AAB31766.1; PID:G688187

A;Accession: I66850

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-75, 'V' <RE2>

A;Cross-references: GB:S73883; NID:G688188

A;Accession: I66851

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-75, 'V' <RE2>

A;Cross-references: GB:S73884; NID:G688189

C;Genetics:

A;Gene: rae-28

C;Superfamily: LAR-interacting protein; SAM homology

F:373-476/Region: glutamine-rich

F:945-1011/Domain: SAM homology <SAM>

Query Match

Best Local Similarity 12.9%; Score 151; DB 2; Length 1012;

Matches 49; Conservative 27; Mismatches 54; Indels 46; Gaps 8;

QY 47 RQFSKAKVKTWPCGP-HQCG-DEGRGAPGSGTRQVFSMAAMNKEGGTASVATGPD 104

DB 864 RGRPRSSSIAKATQGRHGOEDSSRG-----SDNSSYDEALSPTS 906

QY 105 PGFVPLPPG-----KPALPGADGTFPGCPGRKEKP-----SDPVWTVMVDVVEY 149

DB 907 PGFLSVRAGHGERDLGNTITFTPELQGI-----NPVFLSSNFSQWSVEEVEYEF 956

QY 150 FTE-AGFPQATAPQEQEIDGKSLLMQRTDVLTLGSLRGLPALKYEHHIKVLQ 204

DB 957 IASLQCCQBIAEFRSQEIDQALLLKEHLMSAMNKLGPALKICA-KINVLKE 1011

RESULT 2

T03166

probable immediate early protein - alcelaphine herpesvirus 1

C:Species: alcelaphine herpesvirus 1  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C:Accession: T03166  
R:Ensser, A.; Pflanz, R.; Fleckenstein, B.  
J. Virol. 71, 6517-6525, 1997  
A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.  
A:Reference number: Z14840; MUID:97404659; PMID:9261371  
A:Accession: T03166  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1300 <ENS>  
A:Cross-references: UNIPROT:O36421; EMBL:AF005370; NID:G2337967; PIDN:AAC58118.1; PID:G2337967  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 11.6%; Score 136; DB 2; Length 1300;  
Best Local Similarity 31.2%; Pred. No. 0.089;  
Matches 50; Conservative 10; Mismatches 66; Indels 34; Gaps 8;

QY 8 EEEEDDEDEDDVSEG---SEVPESDRPAGAQHHQLNGE--RGQSAKERVKEWTPC 62  
Db 615 EGDGEGPEEPGPGSDPGFGAGQEVPEG--PKGPE-----GECQSGFSSCEGQQVPKGP 667  
QY 63 GPHQ-----QDQEGRGA-PGSGTRQVFSMAAMKEGTTASVATGPDSPSPVLPKPK-- 114  
Db 668 GPEGSSGFGSSEGEFGSGSEGOQVPKAGSSEGEFCRPGGDEDDGDPGPDTEG 727  
QY 115 -----PALPGADGTPFG-----CPPGRKEKPSDPVE 140  
Db 728 EGP CGGPGPDGDGPDGPDGTEGEGCGPGPDGDGDEE 767

RESULT 3  
T13606  
hypothetical protein 87B1.5 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13606; S23632  
R:Murphy, L.; Harris, D.; Barrell, B.  
submitted to the EMBL Data Library, April 1999  
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
A:Reference number: Z17668  
A:Accession: T13606  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1589 <MUR>  
A:Cross-references: UNIPROT:P39769; EMBL:Z98269; NID:e1355202; PID:e1251078; PIDN:CAB109  
R:DeCamillis, M.; Cheng, N.; Pierre, D.; Brock, H.W.  
Genes Dev. 6, 223-232, 1992  
A:Title: The polyhomeotic gene of Drosophila encodes a chromatin protein that shares pol  
A:Reference number: S23632; MUID:92146957; PMID:1346609  
A:Accession: S23632  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1010,'V',1012-1192,'L',1194-1274,'I',1276-1589 <DEC>  
A:Cross-references: EMBL:X63672; NID:g11056; PIDN:CAA45211.1; PID:g11057  
C:Genetics:  
A:Gene: FlyBase:ph-p  
A:Cross-references: FlyBase:FBgn0004861; FlyBase:FBgn0004860  
A:Introns: 12/2; 595/1; 745/2; 1340/1  
C:Superfamily: LAR-interacting protein; SAM homology  
C:Keywords: DNA binding; nucleus  
F:74-80,247-285,411-450,494-650,727-737,775-955,1032-1061/Region: glutamine-rich  
F:1510-1576/Domain: SAM homology <SAM>

Query Match 11.5%; Score 135; DB 2; Length 1589;  
Best Local Similarity 24.0%; Pred. No. 0.13;  
Matches 60; Conservative 38; Mismatches 86; Indels 66; Gaps 12;

QY 2 ERVLEKEEEDDEDE---DEEDDVSEGEVPESDRPAGAQHHQLNGERGQSAKERVKE 58  
Db 1327 QRYADKDVSDPEPPKKATWQEDIKLSGIASAPGSDMVACEQ-----CGKMEHKAKLR 1379  
QY 59 WTPCGP---HQGD---EGRGPAPGSGTRQVFSMAAM-----NKEGGT----- 95

Db 1380 KXYCSFGCSROAKNGIGVGSGETGLTGGIVGVDMALVDRLEDEMAEKKMQTEATPK 1439  
QY 96 -----ASVATGP-----DSPSPVLPKPKP---ALP-----GADGTPFG 126  
Db 1440 LSESFPILGASTEVPMSLPVQAASAFSPPLAMPLGSLPSVALPTLAPLSVVTSGAAPKS 1499  
QY 127 CPPGRKEKPSDPV-BWTYMDVVEYTE-AGPPEQATAFQOEIDGKSLLLMQRTDVLTL 184  
Db 1500 SEVNGTDRP--FISWSVDDVSNFIRELPFGCQDYVDDFTQQEIDGQALLLKEKHLVNAM 1557  
QY 185 SIRLGPALKI 194  
Db 1558 GMLGPAUKI 1567

RESULT 4  
D37271  
A-alpha Z 4 protein - bracket fungus (Schizophyllum commune)  
C:Species: Schizophyllum commune  
C:Date: 07-Feb-1992 #sequence\_revision 07-Feb-1992 #text\_change 09-Jul-2004  
C:Accession: D37271  
R:Ullrich, R.C.  
submitted to the Protein Sequence Database, October 1991  
A:Reference number: A37271  
A:Accession: D37271  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-930 <ULL>  
A:Cross-references: UNIPROT:P37938  
C:Keywords: DNA binding; nucleus; transcription regulation

Query Match 11.1%; Score 130; DB 2; Length 930;  
Best Local Similarity 22.4%; Pred. No. 0.16;  
Matches 55; Conservative 23; Mismatches 84; Indels 84; Gaps 9;

QY 1 ERVLEKEEERDDDEDEDDVSGSEVPSDRPAGAQHHQLNGER-GPQSAKERVKEW 59  
Db 429 DESTDEDDSDSENDSDSEDEEEDDEEEEPVKIAGAKGRNDDDEEVSPLAKPRIFS- 487  
QY 60 TPCGPHQGDGEGRGPAGSGTRQVFSMAAMKEGTTASVATGPDSPSPV----- 108  
Db 488 PPVRPQAIRVSLFSPAPSSR-----GSTPTSPVSPSPKAKRPAQATSL 533  
QY 109 ----PL-----PPKPALPGADGTFP----- 125  
Db 534 ASHPMKREKLEQELRKAGLAPPSAPVLMGPDGVPGLTVRSRSPSPSVSVSLPLP 593  
QY 126 --GCPPGRKEKPSDPVETVMDVVEYFTSAGPPEQATAFQOE-----IDGKSL 174  
Db 594 SRGVFSGGKVTGDTPTPWVND-LEAHTQA--PRDLTAATKSSAGCSVDVPLPGKSRSL 650  
QY 175 MORTDV 180  
Db 651 TRSPSI 656

RESULT 5  
S60254  
nuclear receptor co-repressor N-COR - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
C:Accession: S60254  
R:Hoerlein, A.J.; Naeae, A.M.; Heinzel, T.; Torchia, J.; Gloss, B.; Kurokawa, R.; Ryan,  
Nature 377, 397-404, 1995  
A:Title: Ligand-independent repression by the thyroid hormone receptor mediated by a nuc  
A:Reference number: S60254; MUID:96008539; PMID:7566114  
A:Accession: S60254  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-2453 <HOE>  
A:Cross-references: UNIPROT:Q60974; EMBL:U35312; NID:g1022717; PIDN:AAB17125.1; PID:g1022

```

Query Match      10.6%; Score 124.5; DB 2; Length 2453;
Best Local Similarity 29.2%; Pred. No. 1.1;
Matches 42; Conservative 20; Mismatches 41; Indels 41; Gaps 8;

Qy 2 ERVLKEEEEEDDDEEDDVSGSEVPESDRPAGAQHQLNGERGPGQSAKERVKEWTP 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 523 EKTEKEEEKODKEKDDEDKEDSKETTK--EKDRT-----EATAEPER-EQVTP 569

Qy 62 CGPHQGQDEGRGAPAGSGTRGVFSMAAMNKEGGTASVATGPDSPSPVLPFGPKPALFGAD 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 570 RGRXTANSQGRG--KGRVTR-----SMTSEAAAAAATAATEPPPLPLP----- 612

Qy 122 GTPEGCPCPRKEKPDPVE---WT 142
   || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 613 -----PP--EPISTEPVTSRW 628

RESULT 6
A55817
cyclin-dependent kinase p130-PITSLRE - mouse
C:Species: Mus musculus (house mouse)
C>Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: A55817
R:Malek, S.N.; Desiderio, S.
J. Biol. Chem. 269, 33009-33020, 1994
A>Title: A cyclin-dependent kinase homologue, p130(PITSLRE), is a phosphotyrosine-indepe
A:Reference number: A55817; MUID:95105189; PMID:7528743
C:Suprafamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP
F:424-679/Domain: protein kinase homology <KIN>
F:432-440/Region: protein kinase ATP-binding motif

Query Match      10.6%; Score 123.5; DB 2; Length 783;
Best Local Similarity 26.2%; Pred. No. 0.38;
Matches 64; Conservative 32; Mismatches 85; Indels 63; Gaps 14;

Qy 7 KEHEEDDDDEDDVDVSGSEVPESDRPAGAQQHQ-LNGERGPOS AKERVKEWTPCGPH 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 290 EEEEEESEEGSTSTSESEEEEEEEEEEBEETGSNSEEAESQSAAEVSD-----EE 344

Qy 66 QGDQDEGRG-----PAPGSGTRGVFSMAAMNKEGGTASVATG-----PDSP- 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 MSEDEDENENHILVPESR-----FDRDSGSEGESEVEGETPQSAPTGEDYVPDSPA 400

Qy 106 -SPV-----PLPPGKPAALPGA-----DGTFPGCPGRKEKPSDPVEWTMDVV 147
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 401 LSPIELKQELPYLPALQCSVEEFQCLNRIEGBT-YGVVYRAKKDTDEIVALXELKM 459

Qy 148 EYFTFAAGPQOATAFQEOIIGDKSLLMQRDTDLVTGLSIRLGPAL-KIY-----EHHIK 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 460 EKEKE-GFP--ITSLEI----NTILKAQHPNIIVTVREIIVVGSNMCKIIYVMNVVEHDLK 512

Qy 201 VLQQ 204
   | :
Db 513 SLME 516

RESULT 7
S19560
proline-rich protein MP4 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S19560; S22570
R:Roberts, S.G.E.; Layfield, R.; Bannister, A.J.; McDonald, C.J.
Eur. J. Biochem. 202, 969-974, 1991
A>Title: Gene sequence of mouse B-type proline-rich protein MP4. Transcriptional start p
A:Reference number: S19560; MUID:92111548; PMID:1765104
A:Accession: S19560

```

acidic proline-rich protein precursor - rat  
N;Alternate names: PRP  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 09-Jul-2004  
C;Accession: A03296  
R;Zilemer, M.A.; Swain, W.F.; Rutter, W.J.; Clements, S.; Ann, D.K.; Carlsson, D.M.  
J. Biol. Chem. 259, 10475-10480, 1984  
A;Title: Nucleotide sequence analysis of a proline-rich protein cDNA and peptide homolog  
A;Reference number: A03296; MUID:84289443; PMID:6547951  
A;Accession: A03296  
A;Molecule type: mRNA  
A;Residues: 1-206 <2IE>  
A;Cross-references: UNIPROT:P04474; GB:K02247; NID:G206395; PIDN:AAA41949.1; PID:G206396  
C;Comment: This protein contains six 18- to 19-residue repeats.  
C;Comment: This protein may protect teeth by binding to tannins.  
C;Superfamily: proline-rich protein  
C;Keywords: duplication; parotid gland; saliva; tandem repeat  
F;1-13/Domain: signal sequence #status predicted <SIG>  
F;14-206/Product: acidic proline-rich protein #status predicted <MAT>  
F;80-189/Region: 18-residue repeats  
  
Query Match 10.5%; Score 122.5; DB 1; Length 206;  
Best Local Similarity 26.4%; Pred. No. 0.11;  
Matches 39; Conservative 14; Mismatches 60; Indels 35; Gaps 6;  
  
QY 8 EEEEDDEDEDDVSEGSVEP-----SDRPAGAQHHQLNGERGPOSASAKERVKEW 59  
Db 55 ENGDDGDDSDGDDGDDGNRPERRPPHGGNHQRPDPGHHGPPSGGPGQTSSQGNPQ 114  
  
QY 60 TPCGPHQGQDEG-----RGAPAGSGTRQVFSMAAMNKEGTSVATGDSPPVPLPP 112  
Db 115 GP--PFGGPGQPPQGNPQGPFP-----QGGFQRRPQPGKPGQGP--PQ 156  
  
QY 113 GKPALPGADGTPFGCPP--GRKEKPSDP 138  
Db 157 GGGFPPQPGNPGPPPGGQGHQRRPQP 184  
  
RESULT 10  
B48013  
proline-rich proteoglycan 2 precursor, parotid - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 09-Jul-2004  
C;Accession: B48013  
R;Castle, A.M.; Castle, J.D.  
J. Biol. Chem. 268, 20490-20496, 1993  
A;Title: Novel secretory proline-rich proteoglycans from rat parotid. Cloning and charac  
A;Reference number: A48013; MUID:93388626; PMID:8376404  
A;Accession: B48013  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-295 <CAS>  
A;Cross-references: UNIPROT:Q07611; GB:L17318; NID:G310199; PIDN:AAA03074.1; PID:G310200  
C;Superfamily: proline-rich protein  
C;Keywords: extracellular protein; glycoprotein; tandem repeat  
  
Query Match 10.5%; Score 122.5; DB 2; Length 295;  
Best Local Similarity 26.1%; Pred. No. 0.16;  
Matches 43; Conservative 15; Mismatches 68; Indels 39; Gaps 7;  
  
QY 9 EEEEDDEDEDDVSEGSVPESDRP--GAQHHQ-----LNGERGPOSASAKERVKEW 59  
Db 61 DENGDDNDGDDGDDGDDVNRPERPPQHGGNHHHPHPPAAGPQRPQPSQGGPP-- 119  
  
QY 60 TPCGPHQGQDEG-----RGAPAGSGTRQVFSMAAMNKEGTSVATG--P 102  
Db 120 PFGPQRRPQGGPPPGGQRRPQPGSPQGPFPQGGPQRRPPQGGSP 179  
  
QY 103 DSPSPV-----PLPGKPALPGADGTPFGCPP--GRKEKP 135  
Db 180 QGPPPPGGPQORAPQGPQPPPGGQRRPQPGSPQGPQPPPGGPPQRRP 224

RESULT 11  
H54024  
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE alpha 2-3 - human  
C;Species: Homo sapiens (man)  
C;Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Aug-2004  
C;Accession: H54024; GS4024  
R;Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.  
J. Biol. Chem. 269, 15786-15794, 1994  
A;Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinase  
A;Reference number: A54024; MUID:94253170; PMID:8195233  
A;Accession: H54024  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-768 <XIA>  
A;Cross-references: GB:U04817  
A;Note: authors translated the codon AAA for residue 5 as Leu, GAC for residue 6 as Lys,  
CTC for residue 342 as Phe, CTG for residue 574 as Gln, AAG for residue 614 as Asn, GAT f  
A;Accession: G54024  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-114,117-768 <X12>  
A;Cross-references: GB:U07705  
A;Note: authors translated the codon AAG for residue 5 as Leu, GAC for residue 6 as Lys,  
CTC for residue 351 as Phe, GAT for residue 418 as Glu, AAC for residue 457 as Lys, GTA f  
C;Superfamily: protein kinase homology  
C;Keywords: alternative splicing; ATP; phosphotransferase  
F;409-664/Domain: protein kinase homology <KIN>  
F;417-425/Region: protein kinase ATP-binding motif  
  
Query Match 10.4%; Score 121.5; DB 2; Length 768;  
Best Local Similarity 26.0%; Pred. No. 0.51;  
Matches 63; Conservative 31; Mismatches 83; Indels 65; Gaps 13;  
  
QY 8 EEEEDDEDEDDVSEGSVPESDRPAGAQHHQLNGERGPOSASAKERVKEWTPCGPHQ- 66  
Db 280 EEEEEEEEEEGSTSESESEEEEE---EETGSNSEASESQSAEEVSEEMSEDEER 336  
  
QY 67 -----GQDEGRGAPGSGTRQVFSMAAMNKEGTSVATGPDSP--S 106  
Db 337 ENENHFLVVPESRFRDRDSGESEAEAEVEEGTGPQSALT----EGDYV-----PDSPAL 387  
  
QY 107 PV-----PLPPGKPALPGA-----DGTPEGCPGGRKEKPSDPVETVMDVVEY 149  
Db 388 PIELKQELPKYLPALQGCSEVEEFQCLNRIEST-YGVVYRAKDKYTDIVALKRLKMEK 446  
  
QY 150 FTEAGPPEQATAPQSQEIDGKSLLMQRTDVLTGSLRGPAL-KIY-----EHHIKVL 202  
Db 447 EKE-GFP--ITSIREI-----NTILKAQHENVIVTVREIVVGSNMCKIYIVMNYVEHDLKSL 499  
  
QY 203 QQ 204  
Db 500 ME 501  
  
RESULT 12  
F54024  
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE beta 2-1 - human  
N;Contains: protein kinase cdc2-related PITSLRE beta-1  
C;Species: Homo sapiens (man)  
C;Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Aug-2004  
C;Accession: F54024; D54024  
R;Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.  
J. Biol. Chem. 269, 15786-15794, 1994  
A;Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinase  
A;Reference number: A54024; MUID:94253170; PMID:8195233  
A;Accession: F54024  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-777 <XIA>  
A;Cross-references: GB:U07704  
A;Accession: D54024  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA



Qy 150 FTEAGFPEQATAFQEQRIDGSKLLMQRTDVLTLGSLRGPAL-KIY-----EHHIKVL 202  
 Db 456 EKE-GFP--ITSLEI-----NTILKAQHPNIVTVREIVVGSNMDKIYIVMNVYEHDLKSL 508

Qy 203 QQ 204  
 Db 509 ME 510

RESULT 14  
 E54024  
 protein kinase (EC 2.7.1.37) cdc2-related PITSLRE alpha 2-1 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Aug-2004  
 C/Accession: E54024  
 R/Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.  
 J. Biol. Chem. 269, 15786-15794, 1994  
 A/Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinase  
 A/Reference number: A54024; MUID:94253170; PMID:8195233  
 A/Accession: E54024  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-779 <XIA>  
 A/Cross-references: GB:U04824  
 C/Superfamily: protein kinase homology  
 C/Keywords: alternative splicing; ATP; phosphotransferase  
 F/420-675/Domain: protein kinase homology <KIN>  
 F/428-436/Region: protein kinase ATP-binding motif

Query Match 10.4%; Score 121.5; DB 2; Length 779;  
 Best Local Similarity 26.0%; Pred. No. 0.51;  
 Matches 63; Conservative 31; Mismatches 83; Indels 65; Gaps 13;

Qy 8 EEREDDDEDEEDDVSGSEVPESDRPAGAHQHLNGERGQSAKERVKEWTPCGPHQ- 66  
 Db 291 EEEEEEEEEEGSTSESEEEEE---BEETGSNSEEASEQSAEEVSEBEMSEDEER 347

Qy 67 -----GQDEGRGPAGSGTRQVFSMAAMNKEGTASVATGPDSP--S 106  
 Db 348 ENENHFLVVPESRFDROSGESEAEVEEGTQPSSALT-----EGDYV-----PDSFALS 398

Qy 107 PV-----PLPGKPALPGA-----DGTFFGPPGKKEKPSDPVEVTMDVVEY 149  
 Db 399 PIELKQLPKYLPALQGRSVEEFQCLNRIEETG-YGVVYRAKDKTDEIVALKRLKMEK 457

Qy 150 FTEAGFPEQATAFQEQRIDGSKLLMQRTDVLTLGSLRGPAL-KIY-----EHHIKVL 202  
 Db 458 EKE-GFP--ITSLEI-----NTILKAQHPNIVTVREIVVGSNMDKIYIVMNVYEHDLKSL 510

Qy 203 QQ 204  
 Db 511 ME 512

RESULT 15  
 OZZOMY  
 circumsporozoite protein precursor - Plasmodium yoelii  
 N/Alternate names: sporozoite surface antigen  
 C/Species: Plasmodium yoelii  
 C/Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
 C/Accession: A26271  
 R/Lal, A.A.; de la Cruz, V.F.; Welsh, J.A.; Charoenvit, Y.; Maloy, W.L.; McCutchan, T.F.  
 J. Biol. Chem. 262, 2937-2940, 1987  
 A/Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii  
 A/Reference number: A26271; MUID:87137555; PMID:3102479  
 A/Accession: A26271  
 A/Molecule type: DNA  
 A/Residues: 1-367 <LAL>  
 A/Cross-references: GB:J02695; NID:G150222; PIDN:AAA29558.1; PID:G160223  
 C/Comment: There are three distinct regions in the mature circumsporozoite protein, the  
 obic membrane-anchoring sequence.  
 C/Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repeat.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
C:Keywords: sporozoite; surface antigen; tandem repeat  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-367/Product: circumsporozoite protein #status predicted <MAT>  
F;139-228/Region: 6-residue repeats (Q-G-P-G-A-P)  
F;229-260/Region: 4-residue repeats (Q-Q-P-P)  
F;293-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match	10.2%	Score 119.5;	DB 1;	Length 367;
Best Local Similarity	31.4%	Pred. NO. 0.32;		
Matches	44;	Conservative 13;	Mismatches 50;	Indels 33; Gaps 8;

  

QY	5	LEKEEEEDDEDEDDVSEGESEVPESDRPAGAQHHQHLNGERGQPSAKERVKE--WTPC	62
Db	92	LPKEEKDDLPKEEKDDPPKD---PKDDPPKEAQNKLN---QPVVADENVDOGPGAPQ	145
QY	63	GPHQGDEGRGPAPGSGTRQVFSMAANKEGGTASVATGPDSPS--FVPLPPGKPALPGA	120
Db	146	GP--GAPQGFAPQGFAPQ-----GFGAPQGFAPQGFAPQGFAPQ	185
QY	121	DGTPFGCP--PGRKEKPSDP	138
Db	186	PQGP-GAPQGFAPQGFAPGAP	204

Search completed: September 20, 2005, 12:43:18  
Job time : 25.4357 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 8.92289 Seconds  
(without alignments)  
650.171 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82  
Sequence: 1 EEEEDDEDEDD 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	15	4	AAB82809 Human low
2	82	100.0	26	4	AAB82810 Human low
3	82	100.0	217	2	AAW49041 Human low
4	82	100.0	217	4	AAB82803 Human low
5	82	100.0	241	3	AAB58957 Breast an
6	82	100.0	538	4	AAB82806 Human low
7	75	91.5	358	6	ABB84611 Soybean n
8	73	89.0	457	3	AAY50875 Murine NN
9	73	89.0	1085	2	AAR95607 RRP3 telo
10	73	89.0	1085	7	ADK62464 Disease t
11	72	87.8	114	5	ABG93347 C. albica
12	72	87.8	178	3	AAG29185 Arabidops
13	72	87.8	178	3	AAG34230 Arabidops
14	72	87.8	178	8	ADN74341 Thale cre
15	72	87.8	178	8	ADN73879 Thale cre
16	72	87.8	288	8	ADN24238 Bacterial
17	72	87.8	427	5	ABP73549 Candida a
18	71	86.6	376	5	ADH17117 Human NOV
19	71	86.6	474	5	ADH17115 Human NOV
20	71	86.6	587	8	ADH72068 Human pro
21	71	86.6	589	8	ADRO9338 Human pro
22	71	86.6	591	8	ADH72056 Human pro
23	71	86.6	591	8	ADH72060 Human pro
24	71	86.6	596	8	ADH72058 Human pro
25	71	86.6	597	7	ADD89035 TAT276. 1

26	71	86.6	598	8	ADN02703 Human rec
27	71	86.6	612	5	ADH17114 Human NOV
28	71	86.6	612	6	ABG74699 Human CGD
29	71	86.6	634	8	ADJ67009 Human sec
30	71	86.6	638	6	ABG74700 Human CGD
31	71	86.6	642	8	ADH72066 Human pro
32	71	86.6	646	8	ADH72062 Human pro
33	71	86.6	646	8	ADH72064 Human pro
34	71	86.6	652	6	ABG74695 Human CGD
35	71	86.6	667	5	ADH17113 Human NOV
36	71	86.6	667	8	ADQ18006 Human sof
37	71	86.6	667	8	ABM81423 Tumour-as
38	71	86.6	668	5	ADH16706 Human NOV
39	71	86.6	668	8	ADH72054 Human pro
40	71	86.6	668	8	ADH72072 Human pro
41	71	86.6	668	8	ADH72074 Human pro
42	71	86.6	668	8	ADN42360 Human nov
43	71	86.6	712	7	ADB61623 Rat Prote
44	71	86.6	712	7	ADD46090 Rat Prote
45	71	86.6	712	7	ADE57828 Rat Prote

ALIGNMENTS

RESULT 1  
AAB82809  
ID AAB82809 standard; peptide; 15 AA.  
XX AAB82809;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
XX Homo sapiens.

OS Homo sapiens.

PN WO200164874-A2.

PD 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US006356.

PR 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

(BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

PS Claim 14(a); Page; 143pp; English.

CC The present sequence is that of a peptide comprising amino acid residues 329-343 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. CC Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine

CC compositions. Note: the present sequence is not shown in the  
 CC specification but is derived from the human LBP-2 sequence given in  
 CC figure 7A (see AAB82806)

XX  
 SQ Sequence 15 AA;

Query Match 100.0%; Score 82; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
 |||||  
 Db 1 EEEEDDEDEDEDD 15

RESULT 2  
 ID AAB82810  
 XX AAB82810 standard; peptide; 26 AA.

AC AAB82810;  
 DT 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 2 (LBP-2) peptide.

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX Homo sapiens.

OS  
 PN WO200164874-A2.

XX  
 PD 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

PS Claim 14(a); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues  
 CC 329-354 of novel human low density lipoprotein binding protein 2 (LBP-2,  
 CC see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of  
 CC the invention that are capable of binding to native and methylated low  
 CC density lipoproteins (LDLs). Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs.

CC Polypeptides having the present amino acid sequence are among those  
 CC claimed. Methods of determining if an animal is at risk for  
 CC atherosclerosis, methods for evaluating an agent for use in treating  
 CC atherosclerosis, and methods for treating a cell having an abnormality in  
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
 CC compositions. Note: the present sequence is not shown in the  
 CC specification but is derived from the human LBP-2 sequence given in  
 CC figure 7A (see AAB82806)

XX Sequence 26 AA;

Query Match 100.0%; Score 82; DB 4; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 0.0026;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
 |||||  
 Db 1 EEEEDDEDEDEDD 15

RESULT 3  
 AAW49041

ID AAW49041 standard; protein; 217 AA.

XX AAW49041;

XX 09-NOV-1998 (first entry)

XX Human low density lipoprotein binding protein LBP-2.

XX Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
 KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 8..33 /note= "Claim 2"

FT Peptide 8..32 /note= "Claim 2"

FT Peptide 23..33 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

FT Peptide 208..217 /note= "Claim 2"

QY 1 EEEEDDEDEDEDD 15  
 DB 8 EEEEDDEDEDEDD 22

RESULT 4  
 AAB82803  
 ID AAB82803 standard; protein; 217 AA.  
 XX AAB82803;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Human low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 XX (BOST-) BOSTON HEART FOUND INC.  
 PA  
 XX Lees AM, Lees RS, Law SW, Arjona AA;  
 PI  
 DR WPI: 2001-565505/63.  
 DR N-PSDB; AAB26494.  
 XX

New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

Claim 13(g); Fig 7B; 143pp; English.

The present sequence is that of the N-terminal portion of novel human low density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is deduced from an isolated partial cDNA clone (see AAB26494). A full-length sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins. Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are claimed. Pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine compositions, are also claimed

QY 1 EEEEDDEDEDEDD 15  
 DB 8 EEEEDDEDEDEDD 22

RESULT 5  
 AAB58957  
 ID AAB58957 standard; protein; 241 AA.  
 XX  
 AC AAB58957;  
 XX

Query Match 100.0%; Score 82; DB 4; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 0.022;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
 DB 8 EEEEDDEDEDEDD 22

DT 27-MAR-2001 (first entry)  
 XX  
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.  
 XX  
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
 KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
 KW antidiabetic; antiinflammatory; antitumor; antitumor; antitumor;  
 KW antibacterial; antifungal; antiparasitic; antiparasitic; antiparasitic;  
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
 KW cardiovascular disorder; wound healing; neurological disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055173-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US005881.  
 XX  
 PR 12-MAR-1999; 99US-0124270P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI: 2000-611515/58.  
 DR N-PSDB; AAB21860.  
 XX

New human breast and ovarian cancer associated gene sequences and the polypeptides encoded by these genes, useful in the prevention, treatment and diagnosis of cancer, immune disorders, cardiovascular disorders and neurological diseases.

Claim 11; Page 1112; 1299pp; English.

Sequences AAF21614 - AAF22031 represent DNA sequences encoding human proteins AAB58711 - AAB59128. The DNA and protein sequences are associated with breast and ovarian cancer. Included in the invention are sequences AAF22032 - AAF22040 and AAB59129 which are used in the isolation and characterisation of the DNA and protein sequences of the invention. The breast and ovarian cancer associated DNA, protein, agonist or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; antitumor; antitumor; antitumor; antitumor; antifungal; antiparasitic and cardiant activity. The polynucleotide and protein sequences are used in the diagnosis of cancer, particularly breast and ovarian cancer. The nucleic acid sequences, proteins, agonists and antagonists may also be used in the diagnosis, prevention and treatment of immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular disorders such as myocardial ischaemias; wound healing; neurological diseases such as cerebral anoxia and epilepsy; and infectious diseases

QY 1 EEEEDDEDEDEDD 15  
 DB 32 EEEEDDEDEDEDD 46

RESULT 6  
 AAB82806  
 ID AAB82806 standard; protein; 538 AA.  
 XX  
 AC AAB82806;

Query Match 100.0%; Score 92; DB 3; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
 DB 32 EEEEDDEDEDEDD 46

XX 12-NOV-2001 (first entry)  
 DT Human low density lipoprotein binding protein 2 (LBP-2).  
 DE Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 XX atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 KW Homo sapiens.  
 XX OS  
 XX WO200164874-A2.  
 PN 07-SEP-2001.  
 XX PD  
 XX 28-FEB-2001; 2001WO-US006356.  
 PF 02-MAR-2000; 2000US-00517849.  
 XX PR 14-JUL-2000; 2000US-00616289.  
 XX PA (BOST-) BOSTON HEART FOUND INC.  
 XX PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX WPI; 2001-565505/63.  
 DR N-PSDB; AAH26499.  
 XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX Claim 13(j); Fig 7A; 143pp; English.  
 XX The present sequence is that of novel human low density lipoprotein  
 CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the  
 CC coding region of isolated genomic DNA (see AAH26499). It differs from the  
 CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the  
 CC presence of an additional 321 amino acids at the N-terminus (the cDNA  
 CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP  
 CC polypeptides of the invention that are capable of binding to native and  
 CC methylated low density lipoproteins. Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs.  
 CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529  
 CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of  
 CC determining if an animal is at risk for atherosclerosis, methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed  
 XX Sequence 538 AA;  
 SQ

Query Match 100.0%; Score 82; DB 4; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 0.055;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEEEDDEDEDEDD 15  
 |||||  
 DB 329 EEEEDDEDEDEDD 343

RESULT 7  
 ABB84611  
 ID ABB84611 standard; protein; 358 AA.  
 XX ABB84611;  
 XX 24-FEB-2003 (first entry)  
 DT Soybean nucleosome assembly protein 1.  
 DE NARC10; NARC16; cytosstatic; immunosuppressive; dermatological; cardiant;  
 KW antiinflammatory; nephrotropic; anti-HIV; notropic; neuroprotective;  
 KW antianaemic; cerebroprotective; vasotropic; antidiabetic; anticonvulsant;

KW immunosuppressive; thyromimetic; immunostimulant; antimanic; hypotensive;  
 KW tranquiliser; neuroleptic; gene therapy; gene mapping; apoptosis;  
 KW AIDS; cell cycle disruption; programmed cell death regulation;  
 KW viral infection; nucleosome assembly; phosphate homeostasis;  
 KW cell cycle regulation; cancer; follicular lymphoma; carcinoma;  
 KW p53 mutation; graft rejection; hormone-dependent tumour;  
 KW autoimmune disorder; valvular heart disease;  
 KW systemic lupus erythematosus; diabetes; Hashimoto's thyroiditis;  
 KW immune-mediated glomerulonephritis; virus-induced lymphocyte depletion;  
 KW acquired immunodeficiency syndrome; neurodegenerative disease; stroke;  
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
 KW spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration;  
 KW myelodysplastic syndrome; ischaemic injury; myocardial infarction;  
 KW reperfusion injury; liver disease; idiopathic dilated cardiomyopathy;  
 KW ischaemic cardiomyopathy; aplastic anaemia; chronic neuropania; mania;  
 KW myelodysplastic syndrome; central nervous system disorder; anxiety;  
 KW senile dementia; Huntington's disease; hypertension; schizophrenia;  
 KW severe bipolar affective disorder; nucleosome assembly protein 1.  
 OS Glycine max.  
 XX WO200281516-A2.  
 PN 17-OCT-2002.  
 XX 16-JAN-2002; 2002WO-US001098.  
 PF 16-JAN-2001; 2001US-0262306P.  
 XX PR 15-JAN-2002; 2002US-00047855.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Chiang LW;  
 XX WPI; 2003-058503/05.  
 DR Novel isolated programmed cell death-related polypeptide, NARC10 and  
 XX NARC16, useful for treating disorders associated with abnormal apoptotic  
 PT process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.  
 PT Disclosure; Fig 1A-D; 123pp; English.  
 PS This invention describes novel cell death-related polypeptides NARC10 and  
 CC NARC16, located on chromosome 4q11-4q21 and which have cardiant, antiHIV,  
 CC immunosuppressive, dermatological, antiinflammatory, cerebroprotective,  
 CC notropic, neuroprotective, antianaemic, cardiant, vasotropic, antimanic,  
 CC antidiabetic, immunosuppressive, cytosstatic; thyromimetic, nephrotropic,  
 CC immunostimulant, anticonvulsant, tranquiliser, hypotensive and  
 CC neuroleptic activity and can be used in gene therapy. The products of the  
 CC invention can be used to modulate NARC10 or NARC16 polypeptides or  
 CC polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene  
 CC regions associated with genetic disease or to associate NARC10 or NARC16  
 CC with a disease. The polypeptides are also useful for modulating the  
 CC apoptotic process, and are therefore useful for modulating, and treating  
 CC disorders associated with increased apoptosis, inhibition of apoptosis or  
 CC disruptions in cell cycle, for regulating cellular functions including  
 CC programmed cell death, nucleosome assembly, phosphate homeostasis and the  
 CC cell cycle. Preferably, the products of the invention are useful for  
 CC treating disorders associated with abnormally low rate or abnormally high  
 CC rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas  
 CC with p53 mutations, or hormone-dependent tumours, autoimmune disorders  
 CC including systemic lupus erythematosus, diabetes, graft rejection,  
 CC Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral  
 CC infections e.g. infections caused by herpes viruses, virus-induced  
 CC lymphocyte depletion (including acquired immunodeficiency syndrome  
 CC (AIDS)), neurodegenerative diseases manifested by loss of specific sets  
 CC of neurons (including Alzheimer's disease, Parkinson's disease,  
 CC amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis  
 CC pigmentosa, and cerebellar degeneration), myelodysplastic syndromes  
 CC (including aplastic anaemia), ischaemic injuries (including myocardial  
 CC infarction, stroke and reperfusion injury), and toxin (e.g. alcohol)  
 CC induced liver disease, idiopathic dilated cardiomyopathy, ischaemic  
 CC cardiomyopathy and valvular heart disease, aplastic anaemia, chronic

CC neutropaenia, and myelodysplastic syndromes, central nervous system  
 CC disorders, senile dementia, Huntington's disease, hypertension,  
 CC schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar  
 CC affective disorder (BP-I). This sequence represents the soybean  
 CC nucleosome assembly protein 1 described in the method of the invention  
 XX  
 XX  
 SQ Sequence 358 AA;

Query Match 91.5%; Score 75; DB 6; Length 358;  
 Best Local Similarity 86.7%; Pred. No. 0.24;  
 Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15  
 :||||:|||||  
 DB 307 DEEDEDDEDD 321

RESULT 8  
 ID AAY50875 standard; protein; 457 AA.

XX AAY50875;  
 XX  
 XX 24-FEB-2000 (first entry)  
 XX Murine NNX3 protein.  
 DE NNX3; murine; cytostatic; marker; lung tumor; Hodgkin's disease;  
 XX treatment.  
 KW  
 XX Mus musculus.  
 OS  
 XX WO9960115-A2.  
 PN  
 XX 25-NOV-1999.  
 PD  
 XX 17-MAY-1999; 99WO-EP003374.  
 PF  
 XX 18-MAY-1998; 98EP-00201642.  
 PR  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 PA  
 XX Van Leuven F;  
 PI  
 XX WPI; 2000-053295/04.  
 DR N-PSDB; AAZ43754.  
 XX  
 XX Novel proteins useful as markers and for treating lung tumors and/or  
 PT Hodgkin's disease.  
 PS  
 XX Claim 2; Page 47-49; 50pp; English.  
 XX  
 XX This invention describes two novel NNX3 proteins isolated from human and  
 CC murine sources which have cytostatic activity. The NNX3 mRNA and/or  
 CC protein is useful as a marker for lung tumors or Hodgkin's disease. The  
 CC proteins form pharmaceutical compositions useful for treating lung tumors  
 CC and/or Hodgkin's disease. Polynucleotide products of the invention form  
 CC an assay for screening the expression of these nucleic acids. Antibodies  
 CC raised against the proteins of the invention form an assay for detecting  
 CC the proteins. The probes form an assay for detecting and/or amplifying  
 CC NNX3 polynucleotides. The expression of NNX3 in humans and mice suggest  
 CC that NNX3 is useful as a marker for lung tumors and Hodgkin's disease.  
 CC Therefore, the protein facilitates therapies for these conditions. This  
 CC sequence represents the murine NNX3 protein described in the method of  
 CC the invention  
 XX  
 SQ Sequence 457 AA;

Query Match 89.0%; Score 73; DB 3; Length 457;  
 Best Local Similarity 80.0%; Pred. No. 0.52;  
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15

Db 229 EEEEDDDDDDDDD 243  
 |||||:|:|:|

RESULT 9  
 AAR95607  
 ID AAR95607 standard; protein; 1085 AA.

XX AAR95607;  
 XX  
 XX 12-OCT-1996 (first entry)  
 DT RRP3 telomerase-associated protein.  
 XX  
 XX Yeast; RRP3 protein; telomerase-associated protein; STR7;  
 KW suppressor of telomeric repression-7; telomerase; ribonucleoprotein;  
 KW telomere; tumour; pathogen; sperm; ovum; reporter gene; drug screening;  
 KW antibody; immunoassay; antitumour; antiseptic; contraceptive;  
 KW infertility; diagnostic; gene therapy.  
 XX  
 XX Saccharomyces cerevisiae.  
 OS  
 XX WO9612811-A2.  
 PN  
 XX 02-MAY-1996.  
 PD  
 XX 20-OCT-1995; 95WO-US013801.  
 PF  
 XX 20-OCT-1994; 94US-00326781.  
 PR 28-APR-1995; 95US-00431080.  
 PR  
 XX (ARCH-) ARCH DEV CORP.  
 PA  
 XX Gottschling DE, Singer MS;  
 PI  
 XX WPI; 1996-239169/24.  
 DR N-PSDB; AAT27052.  
 DR  
 XX Novel telomerase associated polypeptide(s) and related nucleic acid -  
 PT useful for detecting e.g. tumour cells or pathogens.  
 XX  
 XX Example 10; Page 294-302; 349pp; English.

XX RRP3 telomerase-associated protein (containing fragment STR7 (AAR95606)  
 CC is found in conjunction with Saccharomyces cerevisiae telomerase, a  
 CC ribonucleoprotein required for telomere replication. Other proteins  
 CC associated with telomerase are given in AAR95601-05 (Other STR proteins).  
 CC These proteins combine with telomerase to repress telomere silencing of  
 CC gene expression. Oligonucleotides from the encoding sequence may be used  
 CC to detect non-ciliate telomerase-associated genes, e.g. in tumour,  
 CC pathogen, sperm or ovum cells. New telomerase-associated sequences may be  
 CC detected by a reporter gene expression system linked to an expression-  
 CC repressing telomerase sequence, and binding compounds, e.g. antibodies, may  
 CC be detected by complex formation with telomerase components. The products  
 CC may be used as antitumour, antiseptic or contraceptive agents, in  
 CC infertility diagnosis, or in gene therapy  
 XX  
 SQ Sequence 1085 AA;

Query Match 89.0%; Score 73; DB 2; Length 1085;  
 Best Local Similarity 80.0%; Pred. No. 1.2;  
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDD 15  
 |||||:|:|:|  
 DB 133 EEEEDDEDDDDDD 147

RESULT 10  
 ADK62464  
 ID ADK62464 standard; protein; 1085 AA.

XX  
 AC ADK62464;

XX 06-MAY-2004 (first entry)  
 DT Disease treating protein complex-derived protein #358.  
 XX protein complex; drug target; diagnosis.  
 XX Unidentified.  
 OS EP1338608-A2.  
 XX 27-AUG-2003.  
 PD 20-DEC-2002; 2002EP-00102902.  
 XX 20-DEC-2001; 2001EP-00130253.  
 PR (CELL-) CELLZONE AG.  
 XX  
 PA Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;  
 PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;  
 PI Michon A, Leutwein C, Rick J;  
 XX WPI; 2003-638460/61.  
 DR N-PSDB; ADK62465.  
 XX New proteins and protein complexes from eukaryotes, useful as targets in  
 PT drug screening, or in diagnosing or screening for the presence of a  
 PT disease or disorder, or a predisposition for developing a disease or  
 PT disorder in a subject.  
 XX Disclosure; SEQ ID NO 715; 13pp; English.  
 PS  
 XX The invention relates to novel protein complexes comprising a first and a  
 CC second protein, or its derivative, fragment, homologue or variant. The  
 CC proteins are selected from given protein complexes, which are not defined  
 CC in the specification. The variants are encoded by nucleic acids that  
 CC hybridize to the nucleic acids encoding the proteins under low stringency  
 CC conditions. The protein complexes are useful as targets for an active  
 CC agent of a pharmaceutical. These protein complexes are particularly  
 CC useful as drugs targets for the treatment or preventing of a disease or  
 CC disorder. The complexes and methods above are useful in diagnosing or  
 CC screening for the presence of a disease or disorder or a predisposition  
 CC for developing a disease or disorder in a subject. These are also useful  
 CC in screening for a drug for treatment or prevention of a disease or  
 CC disorder. The molecule that modulates the amount, activity or protein  
 CC components of the complex is useful for the manufacture of a medicament  
 CC for the treatment or prevention of a disease or disorder. This sequence  
 CC corresponds to a protein of the invention. (Note: the sequence data for  
 CC this patent did not form part of the printed specification but was  
 CC obtained from the EPO in electronic format).  
 XX  
 SQ Sequence 1085 AA;  
 Query Match 89.0%; Score 73; DB 7; Length 1085;  
 Best Local Similarity 80.0%; Pred. No. 1-2;  
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEEEDDEDEDEDD 15  
 Db 133 EEEEDDEDEDEDD 147  
 RESULT 11  
 ID ABG93347  
 XX ABG93347 standard; protein; 114 AA.  
 AC ABG93347;  
 XX 21-NOV-2002 (first entry)  
 DT  
 XX C. albicans BAX-associated protein fragment SEQ ID 652.  
 DE  
 XX

KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;  
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;  
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;  
 KW neurodegeneration; cell death.  
 XX Candida albicans.  
 OS WO200264766-A2.  
 XX 22-AUG-2002.  
 PD 21-DEC-2001; 2001WO-EP015398.  
 XX 22-DEC-2000; 2000EP-00870318.  
 PR 04-JAN-2001; 2000EP-00870002.  
 PR 09-JAN-2001; 2001EP-00870003.  
 XX (JANC ) JANSSEN PHARM NV.  
 PA Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;  
 PI WPI; 2002-667002/71.  
 DR N-PSDB; ABQ76613.  
 XX New isolated nucleic acid representing a synthetic BAX-gene, useful as  
 PT medicament for treating, preventing and/or alleviating yeast or fungal  
 PT infections or proliferative disorders, or for preventing apoptosis in  
 PT certain diseases.  
 XX Claim 36; Fig 2; 344pp; English.  
 PS  
 XX This invention describes a novel nucleic acid representing a synthetic  
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-  
 CC resistant yeast or fungi, identifying, or obtaining and identifying  
 CC Candida spp. sequences that are differentially expressed in a pathway  
 CC eventually leading to programmed cell death or identifying inhibitors or  
 CC inhibitor sequences of Bax-induced cell death. The products of the  
 CC invention have cytostatic, fungicide; immunosuppressive, virucide and  
 CC vasotropic activity and can be used in vaccines or for gene therapy. The  
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,  
 CC antisense molecules and antibodies are useful as medicaments or in  
 CC preparing a medicament for treating, preventing and/or alleviating  
 CC diseases associated with yeast or fungi or proliferative disorders, such  
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds  
 CC or polypeptides, or the genetically modified organism are useful for  
 CC preparing a medicament for modifying the endogenic flora of humans and  
 CC other mammals. The vaccine is useful for immunising against yeast or  
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,  
 CC ischaemia, diseases related with viral infections or neurodegenerations.  
 CC This sequence represents a polypeptide associated with the Bax gene  
 CC described in the disclosure of the invention  
 XX  
 SQ Sequence 114 AA;  
 Query Match 87.8%; Score 72; DB 5; Length 114;  
 Best Local Similarity 80.0%; Pred. No. 0.17;  
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEEEDDEDEDEDD 15  
 Db 22 EEEEDDEDEDEDD 36  
 RESULT 12  
 ID AAG29185  
 XX AAG29185 standard; protein; 178 AA.  
 AC AAG29185;  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 34681.  
 DE  
 XX



KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
PN EPI033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 15-MAY-1999; 99US-0134768P.  
PR 15-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.

PR	01-SEP-1999;	99US-0151930P.	XX	06-SEP-2000.	
PR	07-SEP-1999;	99US-0152363P.	PD		
PR	10-SEP-1999;	99US-0153070P.	XX	25-FEB-2000;	2000EP-00301439.
PR	13-SEP-1999;	99US-0153758P.	PF		
PR	15-SEP-1999;	99US-0154018P.	XX	25-FEB-1999;	99US-0121825P.
PR	16-SEP-1999;	99US-0154039P.	PR	05-MAR-1999;	99US-0123180P.
PR	20-SEP-1999;	99US-0154779P.	PR	09-MAR-1999;	99US-0123548P.
PR	22-SEP-1999;	99US-0155139P.	PR	23-MAR-1999;	99US-0125788P.
PR	23-SEP-1999;	99US-0155486P.	PR	25-MAR-1999;	99US-0126264P.
PR	24-SEP-1999;	99US-0155659P.	PR	29-MAR-1999;	99US-0126785P.
PR	28-SEP-1999;	99US-0156458P.	PR	01-APR-1999;	99US-0127462P.
PR	29-SEP-1999;	99US-0156596P.	PR	06-APR-1999;	99US-0128234P.
PR	04-OCT-1999;	99US-0157117P.	PR	08-APR-1999;	99US-0128714P.
PR	05-OCT-1999;	99US-0157753P.	PR	16-APR-1999;	99US-0129845P.
PR	06-OCT-1999;	99US-0157865P.	PR	19-APR-1999;	99US-0130077P.
PR	07-OCT-1999;	99US-0158029P.	PR	21-APR-1999;	99US-0130449P.
PR	08-OCT-1999;	99US-0158232P.	PR	23-APR-1999;	99US-0130510P.
PR	12-OCT-1999;	99US-0158369P.	PR	23-APR-1999;	99US-0130891P.
PR	13-OCT-1999;	99US-0159293P.	PR	28-APR-1999;	99US-0131449P.
PR	13-OCT-1999;	99US-0159295P.	PR	30-APR-1999;	99US-0132048P.
PR	14-OCT-1999;	99US-0159329P.	PR	30-APR-1999;	99US-0132407P.
PR	14-OCT-1999;	99US-0159330P.	PR	04-MAY-1999;	99US-0132484P.
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAY-1999;	99US-0132485P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-MAY-1999;	99US-0132486P.
PR	14-OCT-1999;	99US-0159638P.	PR	06-MAY-1999;	99US-0132487P.
PR	18-OCT-1999;	99US-0159584P.	PR	07-MAY-1999;	99US-0132863P.
PR	21-OCT-1999;	99US-0160741P.	PR	11-MAY-1999;	99US-0134256P.
PR	21-OCT-1999;	99US-0160767P.	PR	14-MAY-1999;	99US-0134218P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999;	99US-0134219P.
PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999;	99US-0134221P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999;	99US-0134768P.
PR	22-OCT-1999;	99US-0160980P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160981P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999;	99US-0135353P.
PR	22-OCT-1999;	99US-0161404P.	PR	24-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161405P.	PR	25-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161406P.	PR	27-MAY-1999;	99US-0136392P.
PR	26-OCT-1999;	99US-0161359P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161360P.	PR	01-JUN-1999;	99US-0137222P.
PR	26-OCT-1999;	99US-0161361P.	PR	03-JUN-1999;	99US-0137528P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161992P.	PR	07-JUN-1999;	99US-0137724P.
PR	28-OCT-1999;	99US-0161993P.	PR	08-JUN-1999;	99US-0138094P.
PR	29-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138540P.
Query Match 87.8%; Score 72; DB 3; Length 178; Best Local Similarity 80.0%; Pred. No. 0.27; Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	EEEDDEDEDEDD 15			
Db	164	EEEDDDDDDDDEED 178			
: : :					
RESULT 13					
AAG34230					
ID AAG34230 standard; protein; 178 AA.					
XX	AC	AAG34230;			
XX	AC				
DT	18-OCT-2000	(first entry)			
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 41615.				
DE	Protein identification; signal transduction pathway; metabolic pathway;				
XX	hybridization assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
KW	Arabidopsis thaliana.				
XX	EP1033405-A2.				
XX					
PN					

PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 87.8%; Score 72; DB 3; Length 178;  
Best Local Similarity 80.0%; Pred. No. 0.27;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDD 15  
|||||:|:|:|  
Db 164 EEEEDDDDDDEED 178

RESULT 14  
ADN74341  
ID ADN74341 standard; protein; 178 AA.

XX ADN74341;  
AC  
XX  
XX 15-JUL-2004 (first entry)  
DT  
XX  
DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 2236.

XX  
KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;  
KW animal feed product; thale cress; cell wall biosynthesis;  
KW nitrogen metabolism; carbon metabolism.  
XX  
OS Arabidopsis thaliana.

XX WO2004035798-A2.  
PN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 20-OCT-2003; 2003WO-EP011658.  
PF  
XX  
XX 18-OCT-2002; 2002EP-00079408.  
PR  
XX

```
PA (CROP-) CROPDESIGN NV.
XX Inze D, De Veylder L, Vlieghe K;
XX
DR WPI: 2004-348466/32.
DR N-PSDB; ADN73879.
XX
PT Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.
XX
XX Claim 1; SEQ ID NO 2236; 134pp; English.
XX
XX This invention relates to a novel method for altering one or more plant
XX characteristics. Specifically, it refers to identifying genes that are up
XX - or down-regulated in transgenic plants overexpressing the heterodimeric
XX E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX alter plant characteristics accordingly. The present invention describes
XX generating transgenic plants for the production of growth regulators,
XX enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX the altered plant characteristics are selected from increased yield or
XX biomass, enhanced survival capacity, stress tolerance, plant architecture
XX or physiology, altered endoreduplication, biochemistry, signal
XX transduction, storage lipid mobilisation and/or altered photosynthesis,
XX each relative to the corresponding wild type plants. Accordingly, these
XX sequences can also be useful as positive or negative selectable markers
XX during transformation of cells or tissues. The identified genes play a
XX role in a variety of biological processes such as DNA replication, cell
XX wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX transcription factors. This polypeptide sequence is thale cress protein
XX expressed by a gene repressed 1.3 fold or more in plants overexpressing
XX the E2Fa/Dpa transcription factor, given in an exemplification of the
XX invention.
XX
XX SQ Sequence 178 AA;
XX
XX Query Match 87.8%; Score 72; DB 8; Length 178;
XX Best Local Similarity 80.0%; Pred. No. 0.27;
XX Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EEEEDDEDEDEDD 15
XX |||||:|:|:|
XX Db 164 EEEEDDDDDDEED 178
XX
XX RESULT 15
XX ADN73879
XX ID ADN73879 standard; protein; 178 AA.
XX
XX AC ADN73879;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX DE Thale cress protein repressed in E2Fa/Dpa expressing plants SeqID 1774.
XX
XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
XX KW animal feed product; thale cress; cell wall biosynthesis;
XX KW nitrogen metabolism; carbon metabolism.
XX
XX OS Arabidopsis thaliana.
XX
XX PN WO2004035798-A2.
XX
XX PD 29-APR-2004.
XX
XX PF 20-OCT-2003; 2003WO-EP011658.
XX
XX PR 18-OCT-2002; 2002EP-00079408.
XX
XX PA (CROP-) CROPDESIGN NV.
XX
XX PI Inze D, De Veylder L, Vlieghe K;
```

```
XX WPI: 2004-348466/32.
XX N-PSDB; ADN73878.
XX
XX Altering plant characteristics, useful for producing plants for enzyme or
XX pharmaceutical production comprises modifying in a plant, expression of
XX one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
XX
XX Claim 1; SEQ ID NO 1774; 134pp; English.
XX
XX This invention relates to a novel method for altering one or more plant
XX characteristics. Specifically, it refers to identifying genes that are up
XX - or down-regulated in transgenic plants overexpressing the heterodimeric
XX E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX alter plant characteristics accordingly. The present invention describes
XX generating transgenic plants for the production of growth regulators,
XX enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX the altered plant characteristics are selected from increased yield or
XX biomass, enhanced survival capacity, stress tolerance, plant architecture
XX or physiology, altered endoreduplication, biochemistry, signal
XX transduction, storage lipid mobilisation and/or altered photosynthesis,
XX each relative to the corresponding wild type plants. Accordingly, these
XX sequences can also be useful as positive or negative selectable markers
XX during transformation of cells or tissues. The identified genes play a
XX role in a variety of biological processes such as DNA replication, cell
XX wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX transcription factors. This polypeptide sequence is thale cress protein
XX expressed by a gene repressed 1.3 fold or more in plants overexpressing
XX the E2Fa/Dpa transcription factor, given in an exemplification of the
XX invention.
XX
XX SQ Sequence 178 AA;
XX
XX Query Match 87.8%; Score 72; DB 8; Length 178;
XX Best Local Similarity 80.0%; Pred. No. 0.27;
XX Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EEEEDDEDEDEDD 15
XX |||||:|:|:|
XX Db 164 EEEEDDDDDDEED 178
XX
XX Search completed: September 20, 2005, 12:34:42
XX Job time : 10.9229 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 2.27662 Seconds  
(without alignments)  
491.841 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82  
Sequence: 1 BEEDDEDEDEDD 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	3	US-08-979-608A-19
2	82	100.0	15	4	US-09-517-849-19
3	82	100.0	15	4	US-09-616-289-19
4	82	100.0	26	3	US-08-979-608A-20
5	82	100.0	26	4	US-09-517-849-20
6	82	100.0	26	4	US-09-616-289-20
7	82	100.0	217	3	US-08-979-608A-7
8	82	100.0	217	4	US-09-517-849-7
9	82	100.0	217	4	US-09-616-289-7
10	82	100.0	538	4	US-09-616-289-43
11	73	89.0	1085	1	US-08-431-080-28
12	73	89.0	1085	2	US-08-938-534-28
13	73	89.0	1085	3	US-09-345-294-28
14	72	87.8	116	4	US-09-248-796A-17428
15	72	87.8	428	4	US-09-248-796A-15197
16	71	86.6	213	4	US-09-949-016-10813
17	71	86.6	667	4	US-09-538-092-1307
18	71	86.6	1972	4	US-09-418-710-21
19	71	86.6	1972	4	US-09-839-479-21
20	70	85.4	209	4	US-09-214-881A-8
21	70	85.4	214	4	US-09-538-092-883
22	70	85.4	214	4	US-09-214-881A-1
23	70	85.4	226	1	US-08-431-080-26
24	70	85.4	226	2	US-08-938-534-26
25	70	85.4	226	3	US-09-345-294-26
26	70	85.4	706	4	US-09-538-092-957
27	70	85.4	747	4	US-09-949-016-10040

28	70	85.4	932	4	US-09-248-796A-19128	Sequence 19128, A
29	69	84.1	279	4	US-09-699-266A-7	Sequence 7, Appli
30	69	84.1	352	1	US-08-552-142A-4	Sequence 4, Appli
31	69	84.1	352	1	US-08-910-973-4	Sequence 4, Appli
32	69	84.1	352	4	US-09-499-227-4	Sequence 4, Appli
33	69	84.1	352	5	PCT-US95-05741-4	Sequence 4, Appli
34	69	84.1	700	2	US-08-568-459A-10	Sequence 10, Appl
35	69	84.1	700	2	US-08-487-826B-10	Sequence 10, Appl
36	69	84.1	700	3	US-09-210-288-10	Sequence 10, Appl
37	69	84.1	2182	2	US-08-487-826B-16	Sequence 16, Appl
38	68	82.9	764	4	US-09-370-838-67	Sequence 67, Appl
39	68	82.9	764	4	US-09-538-092-944	Sequence 944, App
40	68	82.9	764	4	US-09-854-133-67	Sequence 67, Appl
41	68	82.9	781	4	US-09-949-016-9773	Sequence 9773, Ap
42	68	82.9	1162	2	US-08-728-323A-2	Sequence 2, Appli
43	68	82.9	1162	3	US-09-298-568-2	Sequence 2, Appli
44	68	82.9	1162	4	US-09-410-399-2	Sequence 2, Appli
45	68	82.9	1162	4	US-09-894-273-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-979-608A-19  
; Sequence 19, Application US/08979608A  
; Patent No. 6355451  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/979,608A  
; FILING DATE: 26-NOV-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/048,547  
; FILING DATE: 03-JUN-1997  
; APPLICATION NUMBER: US 60/031,930  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-08-979-608A-19

Query Match 100.0%; Score 82; DB 3; Length 15;

```
Best Local Similarity 100.0%; Pred. No. 0.00065; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 2
US-09-517-849-19
; Sequence 19, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-517-849-19

Query Match 100.0%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 3
US-09-616-289-19
; Sequence 19, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-19

Query Match 100.0%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 4
US-08-979-608A-20
; Sequence 20, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-979-608A-20
Query Match      100.0%; Score 82; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 5
US-09-517-849-20
; Sequence 20, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;                   AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-20

Query Match      100.0%; Score 82; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 7
US-08-979-608A-7
; Sequence 7, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN
;                   DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20
Query Match      100.0%; Score 82; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 6
```

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match      100.0%; Score 82; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDD 15
DB      8 EEEEDDEDEDD 22

RESULT 8
US-09-517-849-7
; Sequence 7, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/517,849
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-7

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match      100.0%; Score 82; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.0083;
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDD 15
DB      8 EEEEDDEDEDD 22

RESULT 9
US-09-616-289-7
; Sequence 7, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-7

Query Match      100.0%; Score 82; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDD 15
DB      8 EEEEDDEDEDD 22

RESULT 10
US-09-616-289-43
; Sequence 43, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
```



```
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-616-289-43

Query Match      100.0%; Score 82; DB 4; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDD 15
    |||||
Db 329 EEEEDDEDEDD 343

RESULT 11
US-08-431-080-28
; Sequence 28, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: US/08/938,534
; FILING DATE:
; APPLICATION NUMBER: 08/431,080
; FILING DATE:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; APPLICATION DATA:
; APPLICATION NUMBER: SN 08/326,781
; FILING DATE: October 20, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-431-080-28

Query Match      89.0%; Score 73; DB 1; Length 1085;
Best Local Similarity 80.0%; Pred. No. 0.39;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDD 15
    |||||
Db 133 EEEEDDEDEDD 147

RESULT 13
US-09-345-294-28
; Sequence 28, Application US/09345294
; Patent No. 6387619
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/345,294
```

```
; FILING DATE: 30-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,080
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-345-294-28

Query Match      89.0%; Score 73; DB 3; Length 1085;
Best Local Similarity 80.0%; Pred. No. 0.39;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDEDD 15
Db      133 EEEEDDEDEDEDD 147

RESULT 14
US-09-248-796A-17428
; Sequence 17428, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17428
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17428

Query Match      87.8%; Score 72; DB 4; Length 116;
Best Local Similarity 80.0%; Pred. No. 0.061;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDEDD 15
Db      24 EEEEDDEDEDEDD 38

RESULT 15
US-09-248-796A-15197
; Sequence 15197, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: 1999-02-12
```

```
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15197
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15197

Query Match      87.8%; Score 72; DB 4; Length 428;
Best Local Similarity 80.0%; Pred. No. 0.21;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 EEEEDDEDEDEDD 15
Db      135 EEEEDDEDEDEDD 149

Search completed: September 20, 2005, 12:45:29
Job time : 3.27662 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 3.94614 Seconds  
(without alignments)  
491.841 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 ESEDDDEDEDDVSEGVSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	26	3	US-08-979-608A-20
2	136	100.0	26	4	US-09-517-849-20
3	136	100.0	26	4	US-09-616-289-20
4	136	100.0	217	3	US-08-979-608A-7
5	136	100.0	217	4	US-09-517-849-7
6	136	100.0	217	4	US-09-616-289-7
7	136	100.0	538	4	US-09-616-289-43
8	104.5	76.8	28	3	US-08-979-608A-26
9	104.5	76.8	28	4	US-09-517-849-26
10	104.5	76.8	28	4	US-09-616-289-26
11	104.5	76.8	232	3	US-08-979-608A-3
12	104.5	76.8	232	4	US-09-517-849-3
13	104.5	76.8	232	4	US-09-616-289-3
14	104.5	76.8	252	3	US-08-979-608A-4
15	104.5	76.8	252	4	US-09-517-849-4
16	104.5	76.8	252	4	US-09-616-289-4
17	104.5	76.8	317	3	US-08-979-608A-2
18	104.5	76.8	317	4	US-09-517-849-2
19	104.5	76.8	317	4	US-09-616-289-2
20	104.5	76.8	550	4	US-09-616-289-47
21	88	64.7	162	1	US-07-879-685B-4
22	88	64.7	599	4	US-09-538-092-864
23	83.5	61.4	305	4	US-09-645-337A-8
24	83	61.0	428	4	US-09-248-796A-15197
25	82	60.3	15	3	US-08-979-608A-19
26	82	60.3	15	4	US-09-517-849-19
27	82	60.3	15	4	US-09-616-289-19

28	81	59.6	557	4	US-09-248-796A-19073	Sequence 19073, A
29	81	59.6	1162	2	US-08-728-323A-2	Sequence 2, Appli
30	81	59.6	1162	3	US-09-298-568-2	Sequence 2, Appli
31	81	59.6	1162	4	US-09-410-399-2	Sequence 2, Appli
32	81	59.6	1162	4	US-09-894-273-2	Sequence 2, Appli
33	80	58.8	706	4	US-09-538-092-957	Sequence 957, App
34	80	58.8	747	4	US-09-949-016-10040	Sequence 10040, A
35	79	58.1	764	1	US-08-375-300-4	Sequence 4, Appli
36	79	58.1	764	3	US-09-177-431-4	Sequence 4, Appli
37	79	58.1	764	5	PCT-US95-16930-4	Sequence 4, Appli
38	79	58.1	1089	1	US-08-375-300-2	Sequence 2, Appli
39	79	58.1	1089	3	US-09-177-431-2	Sequence 2, Appli
40	79	58.1	1089	5	PCT-US95-16930-2	Sequence 2, Appli
41	79	58.1	1972	4	US-09-418-710-21	Sequence 21, Appl
42	79	58.1	1972	4	US-09-839-479-21	Sequence 21, Appl
43	78	57.4	932	4	US-09-248-796A-19128	Sequence 19128, A
44	77	56.6	116	4	US-09-248-796A-17428	Sequence 17428, A
45	76.5	56.2	1085	1	US-08-431-080-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-08-979-608A-20  
; Sequence 20, Application US/08979608A  
; Patent No. 6355451

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq For Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/979,608A

; FILING DATE: 26-Nov. 6355451-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,547

; FILING DATE: 03-JUN-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 26 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-08-979-608A-20

Query Match 100.0%; Score 136; DB 3; Length 26;

```
Best Local Similarity 100.0%; Pred. No. 1.5e-10; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-09-517-849-20
; Sequence 20, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20

Query Match 100.0%; Score 136; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
US-09-616-289-20
; Sequence 20, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-Nov. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-616-289-20

Query Match 100.0%; Score 136; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
US-08-979-608A-7
; Sequence 7, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-Nov. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match 100.0%; Score 136; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 EEEEDDEDEDDVSGSEVPESD 26
| | | | | | | | | | | | | | | | | | | | | | | | | |

US-09-616-289-20
; ORGANISM: Homo sapiens
; TYPE: PRT
; LENGTH: 26
; SEQ ID NO 20
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 53
; PRIOR FILING DATE: 1997-06-03
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 09/517,849
; CURRENT FILING DATE: 2000-07-14
; CURRENT APPLICATION NUMBER: US/09/616,289
; FILE REFERENCE: 10797-004001
; TITLE OF INVENTION: ATHEROSCLEROSIS
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
```

;  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-08-979-608A-7

Query Match 100.0%; Score 136; DB 3; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26  
|||||  
DB 8 EEEEDDEDEDDVSGSEVPESD 33

## RESULT 5

US-09-517-849-7  
; Sequence 7, Application US/09517849  
; Patent No. 6605588  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/517,849  
; FILING DATE: 02-Mar-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/979,608  
; FILING DATE: 26-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-517-849-7

Query Match 100.0%; Score 136; DB 4; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26  
|||||  
DB 8 EEEEDDEDEDDVSGSEVPESD 33

## RESULT 6

US-09-616-289-7  
; Sequence 7, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/616,289  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-616-289-7

Query Match 100.0%; Score 136; DB 4; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26  
|||||  
DB 8 EEEEDDEDEDDVSGSEVPESD 33

## RESULT 7

US-09-616-289-43  
; Sequence 43, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/616,289  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 538  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-616-289-43

Query Match 100.0%; Score 136; DB 4; Length 538;  
Best Local Similarity 100.0%; Pred. No. 3.8e-09;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 EEEEEDEDEDDVSGSEVPESD 26
      |||||:|||||:|||||:|||||
Db      329 EEEEEDEDEDDVSGSEVPESD 354

RESULT 8
US-08-979-608A-26
; Sequence 26, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
;
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1996
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
;
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 28 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-979-608A-26

      Query Match      76.8%; Score 104.5; DB 3; Length 28;
      Best Local Similarity 74.1%; Pred. No. 1.2e-06;
      Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY      1 EEEEEDEDEDDVSGSEVPESD 26
      |||||:|||||:|||||:|||||
Db      2 EEEEEDEDEDDVSGSEVPESD 28

RESULT 10
US-09-616-289-26
; Sequence 26, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
;
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
;
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26

QY      1 EEEEEDEDEDDVSGSEVPESD 26
      |||||:|||||:|||||:|||||
Db      2 EEEEEDEDEDDVSGSEVPESD 28

RESULT 9
US-09-517-849-26
; Sequence 26, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
```

```
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-616-289-26

Query Match          76.8%; Score 104.5; DB 4; Length 28;
Best Local Similarity 74.1%; Pred. No. 1.1e-06;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDD-VSEGSVEPESD 26
   |||:||||:|||||
Db 2 EEEEEDEDDDDVVSEGSVEPESD 28
   |||:||||:|||||

RESULT 11
US-08-979-608A-3
; Sequence 3, Application US/08979608A
; Patent No. 635451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match          76.8%; Score 104.5; DB 3; Length 232;
Best Local Similarity 74.1%; Pred. No. 1.1e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDD-VSEGSVEPESD 26
   |||:||||:|||||
Db 2 EEEEEDEDDDDVVSEGSVEPESD 47
   |||:||||:|||||

RESULT 12
US-08-979-608A-3
; Sequence 3, Application US/08979608A
; Patent No. 635451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-979-608A-3

Query Match          76.8%; Score 104.5; DB 3; Length 232;
Best Local Similarity 74.1%; Pred. No. 1.1e-05;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDD-VSEGSVEPESD 26
   |||:||||:|||||
Db 2 EEEEEDEDDDDVVSEGSVEPESD 47
   |||:||||:|||||

RESULT 13
US-09-616-289-3
; Sequence 3, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
;
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
```

```

Query Match          76.8%; Score 104.5; DB 3; Length 252;
Best Local Similarity 74.1%; Pred. No. 1.2e-05;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 8.28029 Seconds  
(without alignments)  
733.538 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82

Sequence: 1 EEEEDDEDEDD 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 segs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	100.0	15	9	US-09-962-055-19
2	82	100.0	15	9	US-09-976-740-19
3	82	100.0	15	13	US-10-023-523-19
4	82	100.0	15	13	US-10-023-523-19
5	82	100.0	15	15	US-10-616-187-19
6	82	100.0	15	15	US-10-616-187-19
7	82	100.0	26	9	US-09-962-055-20
8	82	100.0	26	9	US-09-976-740-20
9	82	100.0	26	13	US-10-023-523-20
10	82	100.0	26	13	US-10-023-523-20
11	82	100.0	26	15	US-10-616-187-20
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 20, Appl
					Sequence 20, Appl
					Sequence 20, Appl
					Sequence 20, Appl
					Sequence 20, Appl

12	100.0	82	100.0	26	15	US-10-671-242-20	Sequence 20, Appl
13	82	100.0	217	9	US-09-962-055-7	Sequence 7, Appl	
14	82	100.0	217	9	US-09-976-740-7	Sequence 7, Appl	
15	82	100.0	217	13	US-10-023-523-7	Sequence 7, Appl	
16	82	100.0	217	13	US-10-023-523-7	Sequence 7, Appl	
17	82	100.0	217	15	US-10-616-187-7	Sequence 7, Appl	
18	82	100.0	217	15	US-10-671-242-7	Sequence 7, Appl	
19	82	100.0	241	9	US-09-925-298-665	Sequence 665, App	
20	82	100.0	241	14	US-10-103-806-665	Sequence 665, App	
21	82	100.0	538	9	US-09-976-740-43	Sequence 43, Appl	
22	82	100.0	538	13	US-10-023-523-43	Sequence 43, Appl	
23	82	100.0	538	13	US-10-023-523-43	Sequence 43, Appl	
24	82	100.0	538	15	US-10-616-187-43	Sequence 43, Appl	
25	82	100.0	538	15	US-10-671-242-43	Sequence 9, Appl	
26	75	91.5	358	14	US-10-047-855-9	Sequence 172634,	
27	75	91.5	358	15	US-10-424-599-172634	Sequence 40855, A	
28	75	91.5	395	15	US-10-425-114-40855	Sequence 307016,	
29	73	89.0	123	16	US-10-425-115-307016	Sequence 48120, A	
30	72	87.8	80	16	US-10-767-701-48120	Sequence 652, App	
31	72	87.8	114	16	US-10-451-467A-652	Sequence 72908, A	
32	72	87.8	181	15	US-10-425-114-72908	Sequence 6891, Ap	
33	72	87.8	288	15	US-10-369-493-6891	Sequence 7386, Ap	
34	72	87.8	427	14	US-10-032-585-7386	Sequence 169943,	
35	71	86.6	166	15	US-10-424-599-169943	Sequence 9078, Ap	
36	71	86.6	167	16	US-10-739-930-9078	Sequence 653, App	
37	71	86.6	376	15	US-10-072-012-653	Sequence 651, App	
38	71	86.6	474	15	US-10-072-012-651	Sequence 39, Appl	
39	71	86.6	597	15	US-10-331-496A-39	Sequence 650, App	
40	71	86.6	612	15	US-10-072-012-650	Sequence 823, App	
41	71	86.6	667	15	US-10-072-012-649	Sequence 242, App	
42	71	86.6	667	16	US-10-723-860-823	Sequence 21, Appl	
43	71	86.6	668	15	US-10-072-012-242	Sequence 21, Appl	
44	71	86.6	1972	9	US-09-839-479-21	Sequence 21, Appl	
45	71	86.6	1972	15	US-10-376-537-21	Sequence 21, Appl	

ALIGNMENTS

RESULT 1

US-09-962-055-19

; Sequence 19, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; City: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-962-055-19

Query Match      100.0%; Score 82; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 2
US-09-976-740-19
; Sequence 19, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-19

Query Match      100.0%; Score 82; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 3
US-10-023-529-19
; Sequence 19, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-19

Query Match      100.0%; Score 82; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-19

Query Match      100.0%; Score 82; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 4
US-10-023-523-19
; Sequence 19, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-19

Query Match      100.0%; Score 82; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15
```

```
RESULT 5
US-10-616-187-19
; Sequence 19, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-19

Query Match      100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 6
US-10-671-242-19
; Sequence 19, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-19

Query Match      100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 7
US-09-962-055-20
; Sequence 20, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 82; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 8
US-09-976-740-20
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-19

Query Match      100.0%; Score 82; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 7
US-09-962-055-20
; Sequence 20, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 82; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDEDEDEDD 15
Db      1 EEEEDDEDEDEDD 15

RESULT 8
US-09-976-740-20
```

```
; Sequence 20, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-20

Query Match      100.0%; Score 82; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 9
US-10-023-529-20
; Sequence 20, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 82; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 10
US-10-023-523-20
; Sequence 20, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 82; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
DB 1 EEEEDDEDEDEDD 15

RESULT 11
US-10-616-187-20
; Sequence 20, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
```

```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-20

Query Match      100.0%; Score 82; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 12
US-10-671-242-20
; Sequence 20, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671.242
; CURRENT FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-03-02
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-20

Query Match      100.0%; Score 82; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 1 EEEEDDEDEDEDD 15

RESULT 13
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US2002005203A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962.055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match      100.0%; Score 82; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEEEDDEDEDEDD 15
Db 8 EEEEDDEDEDEDD 22

RESULT 14
US-09-976-740-7
; Sequence 7, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-976-740-7

Query Match      100.0%; Score 82; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEEDEDEDEDD 15
Db 8 EEEEEDEDEDEDD 22

RESULT 15
US-10-023-529-7
; Sequence 7, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 82; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEEDEDEDEDD 15
Db 8 EEEEEDEDEDEDD 22

Search completed: September 20, 2005, 12:53:06
Job time : 9.28029 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 7.61934 Seconds  
(without alignments)  
1008.117 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82

Sequence: 1 BEEEDDEDEDEDD 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	100.0	236	2 Q61M4	Q61m4 homo sapien
2	82	100.0	285	2 Q6P0R3	Q6p0r3 homo sapien
3	82	100.0	295	2 Q6P1S7	Q6p1s7 homo sapien
4	82	100.0	538	2 Q6SPF0	Q6spf0 homo sapien
5	79	96.3	239	1 CENB_SHEEP	F49451 ovis aries
6	76	92.7	2246	2 Q8T253	Q8t253 dictyostell
7	75	91.5	358	2 Q3R992	Q3r992 glycine max
8	75	91.5	1260	2 Q7R8Z2	Q7rkz2 plasmodium
9	74	90.2	694	1 NUCLE_CHICK	P15771 gallus gall
10	73	89.0	370	2 Q8BWG3	Q8bwg3 mus musculu
11	73	89.0	409	2 Q8C182	Q8c182 mus musculu
12	73	89.0	457	2 Q9Z243	Q9z243 mus musculu
13	73	89.0	625	2 Q6GNP2	Q6gnp2 xenopus lae
14	73	89.0	1085	1 IFH1_YEAST	P39520 saccharomyc
15	73	89.0	1359	2 Q7S403	Q7s403 neurospora
16	73	89.0	2404	2 Q6L576	Q6l576 plasmodium
17	73	89.0	3933	2 Q97239	Q97239 plasmodium
18	72	87.8	178	2 Q49595	Q49595 arabidopsis
19	72	87.8	215	2 Q81T99	Q81t99 biophalari
20	72	87.8	287	2 Q8I4H9	Q8i4h9 caenorhabdi
21	72	87.8	288	1 SDC_CAREL	P50605 caenorhabdi
22	72	87.8	374	2 Q70Z19	Q70z19 nicotiana t
23	72	87.8	734	2 Q81DC4	Q81dc4 plasmodium
24	72	87.8	785	2 Q9C702	Q9c702 arabidopsis
25	72	87.8	831	2 Q7RIE7	Q7rie7 plasmodium
26	72	87.8	1325	2 Q8IAR4	Q8iar4 plasmodium
27	72	87.8	1978	2 Q7RHX6	Q7rhx6 plasmodium
28	72	87.8	2104	2 Q8IK98	Q8ik98 plasmodium
29	71	86.6	122	2 Q6CQM0	Q6cqmq kluyveromyc
30	71	86.6	123	2 Q6FVV9	Q6fvv9 candida gla
31	71	86.6	126	2 Q75AP4	Q75ap4 ashbya goss

32	71	86.6	137	2 Q6COH4	Q6coh4 yarrowia li
33	71	86.6	198	1 VH23_SFVKA	Q9q907 shope fibro
34	71	86.6	211	1 HMLX_HUMAN	Q9ugv6 homo sapien
35	71	86.6	222	2 Q6PU84	Q6pu84 branchiosto
36	71	86.6	270	2 Q80VC5	Q80vc5 mus musculu
37	71	86.6	373	2 Q9BR60	Q9br60 homo sapien
38	71	86.6	380	2 Q7TNV0	Q7tnv0 mus musculu
39	71	86.6	380	2 Q8BZV6	Q8bzv6 mus musculu
40	71	86.6	393	2 Q7S6W0	Q7s6w0 neurospora
41	71	86.6	612	2 Q61BL2	Q61bl2 homo sapien
42	71	86.6	612	2 Q9BW99	Q9bw99 homo sapien
43	71	86.6	667	1 OS9_HUMAN	Q13438 homo sapien
44	71	86.6	712	1 NUCLE_RAT	P13383 rattus norv
45	71	86.6	727	2 Q60460	Q60460 cricetus

ALIGNMENTS

RESULT 1					
Q96IM4					
ID	Q96IM4	PRELIMINARY;	PRT;	236	AA.
AC	Q96IM4;				
DT	01-DEC-2001 (TREMBlrel. 19, Created)				
DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)				
DT	25-OCT-2004 (TREMBlrel. 28, Last annotation update)				
DE	LOC90378 protein (Hypothetical protein) (Fragment).				
GN	Names=LOC90378;				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Villalon D.K., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Schmutz J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA	Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RA	Strausberg R.;				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Lymph;				
RA	Director MGC Project;				
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC007384; AAH07384.2; -				
DR	EMBL; BC080588; AAH0588.1; -				
DR	HSSP; P39769; 1KW4				
DR	InterPro; IPR001660; SAM				
DR	InterPro; IPR010993; SAM_homology				
DR	Pfam; PF00536; SAM_1; 1				
DR	SMART; SM00454; SAM; 1				

```

DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 82; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 27 EEEEDDEDEDEDD 41

RESULT 2
Q6P0R3 PRELIMINARY; PRT; 285 AA.
ID Q6P0R3
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC065477; AH65477.1; -.
DR InterPro; IPR010660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF1F901 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 76 EEEEDDEDEDEDD 90

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
ID Q6PIS7
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 86 EEEEDDEDEDEDD 100

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
ID Q6SPF0
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 82; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15
Db 86 EEEEDDEDEDEDD 100

```



```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Aorta;
RA  Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL  Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY453840; AAR24087.1; -.
DR  InterPro; IPR001660; SAM.
DR  Pfam; PF00536; SAM_1; 1.
DR  SMART; SM00454; SAM; 1.
DR  PROSITE; PS0105; SAM_DOMAIN; 1.
SQ  SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 82; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 EEEEDDEDEDD 15
Db  329 EEEEDDEDEDD 343

RESULT 5
CENB_SHEEP STANDARD; PRT; 239 AA.
AC P49451;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Major centromere autoantigen B (Centromere protein B) (CENP-B)
DE (Fragment).
GN Name=CENPB;
OS Ovis aries (sheep).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN  [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Burklin D.J., Jones C.A., Burkin H.R., McGrew J.A., Broad T.E.;
RL MEDLINE=97049078; PubMed=8893808;
RT "Sheep CENPB and CENPC genes show a high level of sequence similarity
RT and conserved synteny with their human homologs.";
RL Cytogenet. Cell Genet. 74:86-89(1996).
CC -!- FUNCTION: Interacts with centromeric heterochromatin in
CC chromosomes and binds to a specific subset of alphoid satellite
CC DNA, called the CENP-B box. May organize arrays of centromere
CC satellite DNA into a higher order structure which then directs
CC centromere formation and kinetochore assembly in mammalian
CC chromosomes (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 CENPB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U35655; AAA79098.1; -.
DR InterPro; IPR004875; CENP-B.
KW Centromere; Chromosomal protein; DNA-binding; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 239 AA; 26436 MW; 259C6C72E7D9C135 CRC64;

Query Match 96.3%; Score 79; DB 1; Length 239;
Best Local Similarity 93.3%; Pred. No. 0.85;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 EEEEDDEDEDD 15

```

```

Db  158 EEEEDDEDEDD 172

RESULT 6
Q8T253 Q8T253 PRELIMINARY; PRT; 2246 AA.
AC Q8T253;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN  [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN  [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115683; AAL92693.2; -.
DR InterPro; IPR002110; ANK.
DR SMART; SM00248; ANK; 5.
KW Hypothetical protein.
SQ SEQUENCE 2246 AA; 265604 MW; E99C27A2E19FC67A CRC64;

Query Match 92.7%; Score 76; DB 2; Length 2246;
Best Local Similarity 86.7%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY  1 EEEEDDEDEDD 15
Db  330 EEEEDDEDEDD 344

RESULT 7
Q39892 Q39892 PRELIMINARY; PRT; 358 AA.
AC Q39892;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nucleosome assembly protein 1.
GN Name=SNAP-1;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN  [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Williams; TISSUE=Hypocotyl;
RX MEDLINE=96133687; PubMed=8544812;
RA Yoon H.W., Kim M.C., Lee S.Y., Hwang I., Bahk J.D., Hong J.C.,
RA Ishimi Y., Cho M.J.;
RT "Molecular cloning and functional characterization of a cDNA encoding
RT nucleosome assembly protein 1 (NAP-1) from soybean.";
RL Mol. Gen. Genet. 249:465-473(1995).
CC -!- SIMILARITY: Belongs to the nucleosome assembly protein (NAP)
CC family.
CC EMBL; L38856; AAA88792.1; -.
DR PIR; S60892; S60892.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

```

```
DR GO: 0006334; P:nucleosome assembly; IEA.
DR InterPro; IPR002164; NAP_family.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00956; NAP; 1.
SQ SEQUENCE 358 AA; 40932 MW; C71F9AE9BA04699D CRC64;
Query Match 91.5%; Score 75; DB 2; Length 358;
Best Local Similarity 86.7%; Pred. No. 3.1;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDEDD 15
:||||:|||||
Db 307 DEEDEDEDEDEDD 321

RESULT 8
Q7RKZ2 PRELIMINARY; PRT; 1260 AA.
AC Q7RKZ2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY02756;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perlea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.F., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01000765; EAA22241.1; -.
DR InterPro; IPR008941; TPR-like.
KW Hypothetical protein.
SQ SEQUENCE 1260 AA; 146769 MW; 84BE885A00CF5767 CRC64;
Query Match 91.5%; Score 75; DB 2; Length 1260;
Best Local Similarity 86.7%; Pred. No. 10;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEDEDEDEDEDD 15
|||||:|||||
Db 127 EEEEDDDDDDEDEDE 141

RESULT 9
NUCL_CHICK STANDARD; PRT; 694 AA.
AC P15771;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Nucleolin (Protein C23).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90206792; PubMed=2320420;
RA Maridor G., Nigg E.A.;
RT "cDNA sequences of chicken nucleolin/C23 and N038/B23, two major
RT nucleolar proteins.";
RL Nucleic Acids Res. 18:1286-1286(1990).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=90304215; PubMed=2114180; DOI=10.1016/0167-4781(90)90032-W;
RA Maridor G., Krek W., Nigg E.A.;
RT "Structure and developmental expression of chicken nucleolin and N038:
RT coordinate expression of two abundant non-ribosomal nucleolar
RT proteins.";
RL Biochim. Biophys. Acta 1049:126-133(1990).
RN [3]
RP SEQUENCE OF 407-694 FROM N.A.
RX MEDLINE=89119560; PubMed=2914325; DOI=10.1016/0092-8674(89)90241-9;
RA Borer R.A., Lehner C.F., Eppenberger H.M., Nigg E.A.;
RT "Major nucleolar proteins shuttle between nucleus and cytoplasm.";
RL Cell 56:379-390(1989).
CC -!- FUNCTION: Nucleolin is the major nucleolar protein of growing
CC eukaryotic cells. It is found associated with intranucleolar
CC chromatin and preribosomal particles. It induces chromatin
CC decondensation by binding to histone H1. It is thought to play a
CC role in pre-rRNA transcription and ribosome assembly.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X17199; CAA35080.1; -.
DR EMBL; M21791; AAA48983.1; -.
DR PIR; S08414; DNCNHL.
DR HSSP; P08199; 1FJC.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; RRM_1; 4.
DR SMART; SM00360; RRM; 4.
DR PROSITE; P850102; RRM; 4.
DR DNA-binding; Methylation; Nuclear protein; Phosphorylation; Repeat;
KW RNA-binding.
FT DOMAIN 124 141 Asp/Glu-rich (acidic).
FT DOMAIN 170 192 Asp/Glu-rich (acidic).
FT DOMAIN 217 247 Asp/Glu-rich (acidic).
FT DOMAIN 281 357 RNA-binding (RRM) 1.
FT DOMAIN 371 445 RNA-binding (RRM) 2.
FT DOMAIN 461 535 RNA-binding (RRM) 3.
FT DOMAIN 553 628 RNA-binding (RRM) 4.
FT DOMAIN 630 682 Arg/Gly/Phe-rich.
FT DOMAIN 55 90 5 X 7 AA tandem repeats of X-T-P-X-K-K-X.
FT REPEAT 55 61 1.
FT REPEAT 62 68 2.
FT REPEAT 69 75 3.
FT REPEAT 76 82 4.
FT REPEAT 84 90 5.
FT MOD_RES 116 116 Phosphoserine (By similarity).
FT MOD_RES 136 136 Phosphoserine (By similarity).
FT MOD_RES 171 171 Phosphoserine (By similarity).
FT CONFLICT 419 419 A -> R (in Ref. 3).
FT CONFLICT 520 520 N -> T (in Ref. 3).
SQ SEQUENCE 694 AA; 75640 MW; 7996C504BE9459A1 CRC64;
Query Match 90.2%; Score 74; DB 1; Length 694;
Best Local Similarity 86.7%; Pred. No. 7.2;
Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 1 EBEEDDEDEDEDD 15
Db 173 EBEEDDEDEDEDD 187

RESULT 10
Q8BWG3 PRELIMINARY; PRT; 370 AA.
AC Q8BWG3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 6 day neonate kidney cDNA, RIKEN full-length enriched
DE library, clone:D63000SH08 product:hypothetical Zinc finger, C2H2 type
DE containing protein, full insert sequence.
GN Name=BC035954;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35059500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao K., Nishikawa T., Tashiro H., Itoh M.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Nishine T., Harada A.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Kashiwagi K.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizune W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Konno S., Kondo H., Kouda M., Koya S.,

```

```

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK052616; BAC35064.1; -.
DR MGD; MGI:2673697; BC035954.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW Hypothetical protein
SQ SEQUENCE 370 AA; 40151 MW; EAE89898FD17706 CRC64;
Query Match 89.0%; Score 73; DB 2; Length 370;
Best Local Similarity 86.7%; Pred. No. 5;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EBEEDDEDEDEDD 15
Db 306 EBEEDDEDEDEDD 320
||| |||||:|||||
|| |||||:|||||

RESULT 11
Q8CI82 PRELIMINARY; PRT; 409 AA.
AC Q8CI82;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BC035954 protein.
GN Name=BC035954;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Schmutz J., Smailus D.E., Schnerk A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035954; AAH35954.1; -.
DR MGD; MGI:2673697; BC035954.
DR GO; GO:0005634; C:nucleus; IEA.

```



```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=ATCC 28383 / FL100;
RA  MEDLINE=95304839; PubMed=7785326;
RX  Cherel I., Thuriaux P.;
RA  "The IFH1 gene product interacts with a fork head protein in
RT  Saccharomyces cerevisiae.";
RL  Yeast 11:261-270(1995).
[2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=S288c / AB972;
RA  MEDLINE=973113267; PubMed=9169871;
RX  Johnson M., Hillier L., Riles L., Albermann K., Andre B., Ansoerge W.,
RA  Benes V., Brueckner M., Delius H., Dubois E., Duesterhoft A.,
RA  Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA  Heuss-Netzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA  Louis E.J., Messenguy F., Mewes H.-W., Mioga T., Moestl D.,
RA  Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA  Portetelle D., Furnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA  Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA  Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA  Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA  Wedler H., Zimmermann F.K., Zollner A., Hani J., Hohseil J.D.;
RT  "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL  Nature 387:87-90(1997).
CC  -!- FUNCTION: Controls the pre-rRNA processing machinery in
CC  conjunction with PHL1. Could convert PHL1 from a repressor to an
CC  activator.
CC  -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; Z29488; CAA82624.1; -.
DR  EMBL; U19027; AAB67412.1; -.
DR  PIR; S55352; S55352.
DR  InAC; F39520; -.
DR  Germonline; 142285; -.
DR  SGD; S000004213; IFH1.
DR  GO; GO:0005634; C:nucleus; IDA.
DR  GO; GO:0006348; P:chromatin silencing at telomere; IMP.
DR  GO; GO:0006348; P:RNA processing; IGI.
DR  GO; GO:0006364; P:RNA processing; IGI.
KW  Nuclear protein; Transcription regulation.
FT  DOMAIN 122 163 Asp/Glu-rich (highly acidic).
SQ  SEQUENCE 1085 AA; 122491 MW; BE1C7DEF06213FE0 CRC64;

Query Match      89.0%; Score 73; DB 1; Length 1085;
Best Local Similarity 80.0%; Pred.No. 14;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy  1 EEEEDDDDEDEDD 15
    |||||:|:|
Db  133 EEEEDDDDEDDDD 147

RESULT 15
Q7S403
ID  Q7S403      PRELIMINARY;      PRT; 1359 AA.
AC  Q7S403;
DT  01-MAR-2004 (TrEMBLrel. 26, Created)
DT  01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Hypothetical protein.
GN  Name=NCU05902.1;
OS  Neurospora crassa.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC  Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX  NCBI_TaxID=5141;

```

```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=OR74A;
RA  Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA  Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Renman B.,
RA  Elkins T., Engklev P., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA  Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA  Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA  Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA  Roy A., Foley K., Naylor J., Thoman N., Barrett R., Gnerre S.,
RA  Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frisman D.,
RA  Kryzofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA  Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA  DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA  Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA  Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA  Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT  "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL  Nature 0:0-0(2003).
CC  -!- SIMILARITY: Contains 2 WD repeats.
CC  -!- CAUTION: The sequence shown here is derived from an
CC  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC  preliminary data.
DR  EMBL; AABX01000384; EAA30224.1; -.
DR  InterPro; IPR008938; ARM.
DR  InterPro; IPR000547; Clathrin_repeat.
DR  InterPro; IPR001199; Cyt_B5.
DR  InterPro; IPR001680; WD40.
DR  InterPro; IPR011046; WD40_like.
DR  Pfam; PF00637; Clathrin; 1.
DR  Pfam; PF00400; WD40; 2.
DR  PROSITE; PS00191; CYTOCHROME B5_1; UNKNOWN_1.
KW  Hypothetical protein; Repeat; WD repeat.
SQ  SEQUENCE 1359 AA; 149201 MW; BDCB61B70764D641 CRC64;

Query Match      89.0%; Score 73; DB 2; Length 1359;
Best Local Similarity 86.7%; Pred.No. 17;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy  1 EEEEDDDDEDEDD 15
    |||||:|:|
Db  76 ETEDEDDDEDEDD 90

Search completed: September 20, 2005, 12:41:43
Job time : 9.61934 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 1.68911 Seconds  
(without alignments)  
854.447 Million cell updates/sec

Title: US-10-671-242-19

Perfect score: 82  
Sequence: 1 EEEDEDEDEDD 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	91.5	358	2 S60892	nucleosome assembl
2	74	90.2	694	1 DNCNHL	nucleolin - chicke
3	73	89.0	1085	2 S55352	IFH1 protein - yea
4	72	87.8	178	2 T51159	HMG protein (impor
5	72	87.8	288	2 T22846	hypothetical prote
6	71	86.6	667	2 JC5889	OS-9 protein precu
7	71	86.6	712	2 JH0148	nucleolin - rat
8	71	86.6	727	2 JC5113	ribosomal transcri
9	71	86.6	764	2 JC5112	ribosomal transcri
10	70	85.4	215	2 S02826	nonhistone chromos
11	70	85.4	562	2 S38149	SIS2 protein - yea
12	70	85.4	707	2 A35804	nucleolin - human
13	70	85.4	802	1 S48529	NAB3 protein - yea
14	70	85.4	1105	2 T18295	Ap-3 adaptor compl
15	69	84.1	279	2 S30766	ASF1 protein - yea
16	69	84.1	352	2 T51687	neurogenic differe
17	69	84.1	636	2 T51893	related to Che-1 p
18	69	84.1	867	2 T40336	probable MGF trans
19	69	84.1	2182	2 T28634	variant-specific s
20	68	82.9	379	2 C84577	probable nucleosom
21	68	82.9	389	2 T27574	hypothetical prote
22	68	82.9	565	2 T07424	phosphonositide-s
23	68	82.9	678	2 A54514	glutamic acid-rich
24	68	82.9	700	2 S38426	chaperonin 60 prec
25	68	82.9	727	2 S18193	autoantigen NOR-90
26	68	82.9	764	2 S09318	transcription fact
27	68	82.9	792	2 T42963	hypothetical prote
28	68	82.9	896	2 T24169	hypothetical prote
29	68	82.9	906	2 T24166	hypothetical prote

ALIGNMENTS

RESULT 1

S60892

nucleosome assembly protein 1 - soybean

C;Species: Glycine max (soybean)

C;Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S60892

R;Yoon, H.W.; Kim, M.C.; Lee, S.Y.; Hwang, I.; Bahk, J.D.; Hong, J.C.; Ishimi, Y.; Cho, I.

Mol. Gen. Genet. 249: 465-473, 1995

A;Title: Molecular cloning and functional characterization of a cDNA encoding nucleosome

A;Reference number: S60892; MUID:96133687; PMID:8544812

A;Accession: S60892

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-358 <YOO>

A;Cross-references: UNIPROT:Q39892; EMBL:L38856; NID:g1161251; PIDN:AAA88792.1; PID:g1161

Query Match 91.5%; Score 75; DB 2; Length 358;

Best Local Similarity 86.7%; Pred. No. 0.28;

Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EEEDEDEDEDD 15

Db 307 DEEDEDDEDD 321

RESULT 2

DNCNHL

nucleolin - chicken

N;Alternate names: nucleolar protein C23

C;Species: Gallus gallus (chicken)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C;Accession: S08414; S10766; A32725; I50397; B30099

R;Maridor, G.; Nigg, E.A.

Nucleic Acids Res. 18, 1286, 1990

A;Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar proteins

A;Reference number: S08414; MUID:90206792; PMID:2320420

A;Accession: S08414

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-694 <WAR>

A;Cross-references: UNIPROT:P15771; EMBL:X17199; NID:g63710; PIDN:CAA35060.1; PID:g63711

R;Maridor, G.; Krek, W.; Nigg, E.A.

Biochim. Biophys. Acta 1049, 126-133, 1990

A;Title: Structure and developmental expression of chicken nucleolin and NO38: coordinates

A;Reference number: S10766; MUID:90304215; PMID:2114180

A;Accession: S10766

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-694 <WAR>

R;Peters, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.

Cell 60, 791-801, 1990

A;Title: Identification of major nucleolar proteins as candidate mitotic substrates of c

```

Qy      1 EEEEDDDDEDEDD 15
      |||||:|:|
Db      133 EEEEDDDDEDDDD 147

RESULT 4
T51159      HMG protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 16-Aug-2004
C/Accession: T51159; T51167
R/Comella, P.; Wu, H.J.; Laudie, M.; Berger, C.; Cooke, R.; Delseny, M.; Grellier,
Plant Mol. Biol. 41, 687-700, 1999
A/Title: Fine sequence analysis of 60 kb around the Arabidopsis thaliana AtEm1
A/Reference number: Z24835; MUID:20108326; PMID:10645728
A/Accession: T51159
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-178 <COM>
A/Cross-references: UNIPROT:O49595; EMBL:AF049236; PIDN:AAC14415.1
R/Stemmer, C.; Ritt, C.; Igloi, G.L.; Grimm, R.; Grasser, K.D.
Eur. J. Biochem. 250, 646-652, 1997
A/Title: Variability in Arabidopsis thaliana chromosomal high-mobility-group-1-
A/Reference number: Z13511; MUID:98121184; PMID:9461286
A/Accession: T51167
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-178 <STE>
A/Cross-references: EMBL:Y14071; PIDN:CAA74400.1
A/Experimental source: strain DK105
C/Genetics:
A/Gene: HMGalpha
A/Map position: 3
A/Introns: 23/2; 64/2; 83/3; 98/3; 125/3; 148/3
C/Superfamily: HMG box homology

Query Match      87.8%; Score 72; DB 2; Length 178;
Best Local Similarity 80.0%; Pred. No. 0.29;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDDDEDEDD 15
      |||||:|:|
Db      164 EEEEDDDDDDEED 178

RESULT 5
T22846      hypothetical protein F57C7.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T22846
R/White, S.
submitted to the EMBL Data Library, February 1996
A/Reference number: Z19625
A/Accession: T22846
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-288 <WIL>
A/Cross-references: UNIPROT:P50605; EMBL:Z69646; PIDN:CAA93474.1; GSPDB:GN000028
A/Experimental source: clone F57C7
C/Genetics:
A/Gene: CRSP:F57C7.3
A/Map position: X
A/Introns: 51/3; 87/1; 120/3; 191/2; 231/1

Query Match      87.8%; Score 72; DB 2; Length 288;
Best Local Similarity 80.0%; Pred. No. 0.47;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EEEEDDDDEDEDD 15
      |||||:|:|
Db      142 EEEEDDDDEDEDE 156

```



RESULT 6  
JCS989  
OS-9 protein precursor - human  
N:Contains: OS-9 protein, splice form 1; OS-9 protein, splice form 2; OS-9 protein, splice form 3  
C:Species: Homo sapiens (man)  
C>Date: 18-Mar-1998 #sequence\_revision 18-Mar-1998 #text\_change 09-Jul-2004  
C:Accession: JCS989; JEO106; JEO107; JEO108  
R:Kimura, Y.; Nakazawa, M.; Tsuchiya, N.; Asakawa, S.; Shimizu, N.; Yamada, M.  
J. Biochem. 122, 1190-1195, 1997  
A:Title: Genomic organization of the OS-9 gene amplified in human sarcomas.  
A:Reference number: JCS989; MUID:98158329; PMID:9498564  
A:Accession: JCS989  
A:Molecule type: DNA  
A:Residues: 1-667 <KIM1>  
A:Cross-references: UNIPROT:Q13438; DDBJ:AB002806; NID:G2780782; PIDN:BAA24363.1; PID:d10252775  
A:Experimental source: sarcomas  
A>Note: neither the complete nucleic acid sequence nor the complete translation are shown  
R:Kimura, Y.; Nakazawa, M.; Yamada, M.  
J. Biochem. 123, 876-882, 1998  
A:Title: Cloning and characterization of three isoforms of OS-9 cDNA and expression of the OS-9 protein  
A:Reference number: JEO108; MUID:98230694; PMID:9562620  
A:Accession: JEO106  
A:Molecule type: mRNA  
A:Residues: 1-667 <KIM2>  
A:Cross-references: DDBJ:AB002806; NID:G2780782; PIDN:BAA24363.1; PID:d10252775; PID:G2780782  
A:Accession: JEO107  
A:Molecule type: mRNA  
A:Residues: 1-534,590-667 <KIM3>  
A:Cross-references: DDBJ:AB002806  
A:Accession: JEO108  
A:Molecule type: mRNA  
A:Residues: 1-455,471-534,590-667 <KIM4>  
A:Cross-references: DDBJ:AB002806  
C:Comment: This protein is involved in amplification and overexpression of various tumors  
C:Genetics:  
A:Gene: OS-9  
A:Cross-references: GDB:9958646  
A:Map position: 12q13-12q15  
C:Keywords: alternative splicing; carcinogenesis; glycoprotein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:46-667/Product: OS-9 protein, splice form 1 #status predicted <MAT1>  
F:26-534,590-667/Product: OS-9 protein, splice form 2 #status predicted <MAT2>  
F:26-455,471-534,590-667/Product: OS-9 protein, splice form 3 #status predicted <MAT3>  
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 86.6%; Score 71; DB 2; Length 667;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
|||||:|||||:  
Db 414 EEEEDDEDEDEDD 428

RESULT 7  
JH0148  
nucleolin - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 23-Jul-1999  
C:Accession: JH0148; A24088; I63130  
R:Bourbon, H.M.; Amalric, F.  
Gene 88, 187-196, 1990  
A:Title: Nucleolin gene organization in rodents: highly conserved sequences within three exons  
A:Reference number: JH0148; MUID:90269607; PMID:2347493  
A:Accession: JH0148  
A:Molecule type: DNA  
A:Residues: 1-712 <BOU>  
A>Note: the authors translated the initiation codon GTG for residue 1 as Val  
R:Lischwe, M.A.; Cook, R.G.; Ahn, Y.S.; Yeoman, L.C.; Busch, H.  
Biochemistry 24, 6025-6028, 1985  
A:Title: Clustering of glycine and Ng,Ng-dimethylarginine in nucleolar protein C23.  
A:Reference number: A24088; MUID:86104094; PMID:4084504

A:Accession: A24088  
A:Molecule type: protein  
A:Residues: 651-703 <LIS>  
R:Bourbon, H.  
Gene 89, 73-84, 1988  
A:Title: Sequence and structure of the nucleolin promoter in rodents: Characterization of the promoter in hamster and rat  
A:Reference number: I48118; MUID:89121496; PMID:2906027  
A:Accession: I63130  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 'MV', 2-44 <RES>  
A:Cross-references: GB:M2090; NID:G205793; PIDN:AAA1733.1; PID:G205794  
C:Comment: This protein is the major nucleolar-specific protein in eukaryotic exponential phase cells  
C:Genetics:  
A:Gene: nuc  
A:Start codon: GTG  
A:Introns: 5/3; 44/3; 211/1; 274/1; 303/1; 350/2; 392/1; 433/2; 485/1; 526/2; 571/1; 613/1  
C:Superfamily: nucleolin; ribonucleoprotein repeat homology  
C:Keywords: DNA binding; nucleus  
F:311-376/Domain: ribonucleoprotein repeat homology <RRM1>  
F:397-459/Domain: ribonucleoprotein repeat homology <RRM2>  
F:489-552/Domain: ribonucleoprotein repeat homology <RRM3>  
F:575-639/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 86.6%; Score 71; DB 2; Length 712;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
|||||:|||||:  
Db 241 EEEEDDEDEDEDD 255

RESULT 8  
JCS113  
ribosomal transcription factor UBF2 - Chinese hamster  
C:Species: Crictetus griseus (Chinese hamster)  
C>Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 16-Aug-2004  
C:Accession: JCS113  
R:Bollivar, J.; Goenechea, L.G.; Grenett, H.; Pendon, C.; Valdivia, M.M.  
Gene 176, 257-258, 1996  
A:Title: Cloning and sequencing of the genes encoding the hamster ribosomal transcription factor UBF2  
A:Reference number: JCS112; MUID:97075939; PMID:8918262  
A:Accession: JCS113  
A:Molecule type: mRNA  
A:Residues: 1-727 <BOL>  
A:Cross-references: UNIPROT:Q60460; GB:I42571  
C:Comment: This factor binds to the ribosomal RNA gene's promoter and forms a stable pre-initiation complex  
C:Superfamily: HMG box homology  
F:109-184/Domain: HMG box homology <HMG1>  
F:257-329/Domain: HMG box homology <HMG3>  
F:445-516/Domain: HMG box homology <HMG5>

Query Match 86.6%; Score 71; DB 2; Length 727;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
|||||:|||||:  
Db 646 EEEEDDEDEDEDD 660

RESULT 9  
JCS112  
ribosomal transcription factor UBF1 - Chinese hamster  
C:Species: Crictetus griseus (Chinese hamster)  
C>Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 16-Aug-2004  
C:Accession: JCS112  
R:Bollivar, J.; Goenechea, L.G.; Grenett, H.; Pendon, C.; Valdivia, M.M.  
Gene 176, 257-258, 1996  
A:Title: Cloning and sequencing of the genes encoding the hamster ribosomal transcription factor UBF1  
A:Reference number: JCS112; MUID:97075939; PMID:8918262  
A:Accession: JCS112

A:Molecule type: mRNA  
A:Residues: 1-764 <BOL>  
A:Cross-references: UNIPROT:Q60459; GB:I42570  
C;Comment: This factor binds to the ribosomal RNA gene's promoter and forms a stable pre-  
C;Superfamily: HMG box homology  
F;109-184/Domain: HMG box homology <HMG1>  
F;193-268/Domain: HMG box homology <HMG2>  
F;294-366/Domain: HMG box homology <HMG3>

Query Match 86.4%; Score 71; DB 2; Length 764;  
Best Local Similarity 80.0%; Pred. No. 1.5;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEEDD 15  
||:||||:|:||||  
Db 683 EEEEDDEDEEDD 697

RESULT 10  
S02826  
nonhistone chromosomal protein HMG-1 - human  
C;Species: Homo sapiens (man)  
C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text\_change 09-Jul-2004  
C;Accession: S02826; A33178; G33178  
R;Wen, L.; Huang, J.K.; Johnson, B.H.; Reeck, G.R.  
Nucleic Acids Res. 17, 1197-1214, 1989  
A;Title: A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-1.  
A;Reference number: S02826; MUID:89160247; PMID:2922262  
A;Accession: S02826  
A:Molecule type: mRNA  
A:Residues: 1-215 <WEN>  
A:Cross-references: UNIPROT:P09429; EMBL:X12597; NID:G32326; PIDN:CAA31110.1; PID:G32327  
R;Ward, L.D.; Hong, J.; Whitehead, R.H.; Simpson, R.J.  
Electrophoresis 11, 883-891, 1990  
A;Title: Development of a database of amino acid sequences for human colon carcinoma pro  
A;Reference number: A33178; MUID:91176935; PMID:2079031  
A;Accession: A33178  
A:Molecule type: protein  
A:Residues: 2-13, 'XXF' <WAR>  
A;Accession: G33178  
A:Molecule type: protein  
A:Residues: 2-13, 'XX', 16-22 <WA2>  
C;Genetics:  
A;Gene: GDB:HWG1  
A;Cross-references: GDB:133789; OMIM:163905  
A;Map position: 13q12-13q12  
C;Superfamily: nonhistone chromosomal protein HMG-2; HMG box homology  
C;Keywords: chromosomal protein; DNA binding; nucleus  
F;6-83/Domain: HMG box homology <HMG1>  
F;92-166/Domain: HMG box homology <HMG2>

Query Match 85.4%; Score 70; DB 2; Length 215;  
Best Local Similarity 80.0%; Pred. No. 0.56;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEEDD 15  
||:||||:|:||||  
Db 197 EEEEDDEDEEDD 211

RESULT 11  
S38149  
S1S2 protein - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein YKR072C  
C;Species: Saccharomyces cerevisiae  
C;Date: 03-May-1994 #sequence revision 03-May-1994 #text\_change 09-Jul-2004  
C;Accession: S38149; S54982; S43078  
R;Pohl, T.M.; Pohl, F.M.  
submitted to the Protein Sequence Database, March 1994  
A;Reference number: S37897  
A;Accession: S38149  
A:Molecule type: DNA  
A:Residues: 1-562 <POH>

A:Cross-references: UNIPROT:P36024; EMBL:Z28297; NID:g486544; PID:g486545; MIPS:YKR072C  
A;Experimental source: strain S288C  
R;di Como, C.J.; Bose, R.; Arndt, K.T.  
Genetics 139, 95-107, 1995  
A;Title: Overexpression of S1S2, which contains an extremely acidic region, increases the  
F;109-184/Domain: HMG box homology <HMG1>  
A;Reference number: S54982; MUID:95220693; PMID:7705654  
A;Accession: S54982  
A:Molecule type: DNA  
A:Residues: 1-562 <DIW>  
A:Cross-references: EMBL:U01878; NID:g430983; PIDN:AAA80000.1; PID:g430984  
C;Genetics:  
A;Gene: SGD:S1S2  
A:Cross-references: SGD:S0001780; MIPS:YKR072C  
A;Map position: 11R  
C;Keywords: nucleus

Query Match 85.4%; Score 70; DB 2; Length 562;  
Best Local Similarity 85.7%; Pred. No. 1.4;  
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EEEEDDEDEEDD 15  
||:||||:|:||||  
Db 496 EEEEDDEDEEDD 509

RESULT 12  
A35804  
nucleolin - human  
N;Alternate names: phosphoprotein pp100; protein B50; protein C23  
C;Species: Homo sapiens (man)  
C;Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text\_change 09-Jul-2004  
C;Accession: A35804; S04631; A48138; A55996  
R;Srivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.  
J. Biol. Chem. 265, 14922-14931, 1990  
A;Title: Genomic organization and chromosomal localization of the human nucleolin gene.  
A;Reference number: A35804; MUID:90368666; PMID:2394707  
A;Accession: A35804  
A:Molecule type: DNA  
A:Residues: 1-707 <SR1>  
A:Cross-references: UNIPROT:P19338; GB:M60858; GB:J05584; NID:g189305; PIDN:AAA59954.1; I  
R;Srivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.  
FEBS Lett. 250, 99-105, 1989  
A;Title: Cloning and sequencing of the human nucleolin cDNA.  
A;Reference number: S04631; MUID:89290043; PMID:2737305  
A;Accession: S04631  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-707 <SR2>  
R;Tshikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.  
Mol. Cell. Biol. 13, 4301-4310, 1993  
A;Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and t  
A;Reference number: A48138; MUID:93309464; PMID:8321232  
A;Accession: A48138  
A:Molecule type: protein  
A:Residues: 458-474 <ISH>  
A;Experimental source: HeLa cell nuclei  
A;Note: sequence extracted from NCBI backbone (NCBIP:134645)  
R;Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.  
Biochemistry 33, 14696-14706, 1994  
A;Title: Major cell surface-located protein substrates of an ecto-protein kinase are hom  
A;Reference number: A55996; MUID:95086063; PMID:7993898  
A;Accession: A55996  
A:Molecule type: protein  
A:Residues: 231-236;349-362;399-403;458-461;655-656, 'X', 658-660 <JOR>  
A;Experimental source: surface-labelled HeLa cells  
C;Genetics:  
A;Gene: GDB:NCL  
A;Cross-references: GDB:125908; OMIM:164035  
A;Map position: 2q12-2qter  
C;Superfamily: nucleolin; ribonucleoprotein repeat homology  
C;Keywords: DNA binding; nucleolus; nucleus; phosphoprotein; RNA binding  
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>  
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>

F:487-550/Domain: ribonucleoprotein repeat homology <RRM3>  
F:573-634/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 85.4%; Score 70; DB 2; Length 707;  
Best Local Similarity 80.0%; Pred. No. 1.8;  
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
:|||||:|||||  
Db 146 DEEEDDDSEEDDD 160

## RESULT 13

S48529  
NAB3 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: probable RNA/asDNA-binding protein HMD1; protein P1945; protein YPL19  
C:Species: Saccharomyces cerevisiae  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S48529; S65209; S60122  
R:Wilson, S.M.; Oberdorf, A.M.; Datar, K.V.; Swedlow, J.R.; Paddy, M.R.; Swanson, M.S.  
submitted to the EMBL Data Library, January 1994  
A:Description: Characterization of Nuclear Polyadenylated RNA-Binding Proteins from Sac  
A:Reference number: S48529  
A:Accession: S48529  
A:Molecule type: DNA  
A:Residues: 1-802 <WIL>  
A:Cross-references: UNIPROT:P38996; EMBL:U05314; NID:g476219; PID:g476220  
R:Rieger, M.; Mueller-Auer, S.; Schaefer, M.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S65202  
A:Accession: S65209  
A:Molecule type: DNA  
A:Residues: 1-802 <RIE>  
A:Cross-references: EMBL:Z73546; NID:g1370396; PID:g1370397; GSPDB:GN00016; MIPS:YPL190  
A:Experimental source: strain S288C (AB372)  
R:Sugimoto, K.; Matsumoto, K.; Kornberg, R.D.; Reed, S.I.; Wittenberg, C.  
Mol. Gen. Genet. 248, 712-718, 1995  
A:Title: Dosage suppressors of the dominant G1 cyclin mutant CLN3-2: identification of a  
A:Reference number: S60122; MUID:96069710; PMID:7476874  
A:Accession: S60122  
A>Status: nucleic acid sequence not shown

A:Molecule type: DNA  
A:Residues: 1-340,'I',342-802 <SUG>  
A:Cross-references: GB:D37935; NID:g1235749; PID:g1235750  
C:Genetics:  
A:Gene: SGD:NAB3; NAB3; HMD1; MIPS:YPL190C  
A:Cross-references: SGD:S000611  
A:Map position: 16L  
C:Superfamily: yeast NAB3 protein; ribonucleoprotein repeat homology  
C:Keywords: nucleus; RNA binding  
F:331-396/Domain: ribonucleoprotein repeat homology <RRM1>

Query Match 85.4%; Score 70; DB 1; Length 802;  
Best Local Similarity 73.3%; Pred. No. 2;  
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
:|||||:|||||  
Db 112 EEEEDDDDDDDDD 126

## RESULT 14

T18295  
Ap-3 adaptor complex beta3A chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18295  
R:Feng, L.; Seymour, A.B.; Jiang, S.Y.; To, A.; Peden, A.A.; Novak, E.K.; Zhen, L.; Rusi  
Hum. Mol. Genet. 8, 323-330, 1999  
A:Title: The beta3A subunit gene (Ap3b1) of the AP-3 adaptor complex is altered in the m  
A:Reference number: Z18864; MUID:99135912; PMID:9931340  
A:Accession: T18295  
A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA  
A:Residues: 1-1105 <FEN>  
A:Cross-references: UNIPROT:O921T1; EMBL:AF103809; NID:g3885987; PID:g3885988; PIDN:AAC7  
A:Experimental source: strain C3H/HeJ  
C:Genetics:  
A:Gene: Ap3b1  
A:Map position: 13

Query Match 85.4%; Score 70; DB 2; Length 1105;  
Best Local Similarity 80.0%; Pred. No. 2.7;  
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
:|||||:|||||  
Db 681 EEEEDDEDEDEEE 695

## RESULT 15

S30766  
ASF1 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein J0755; protein YJL115W  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-May-1993 #sequence\_revision 28-May-1993 #text\_change 09-Jul-2004  
C:Accession: S30766; S56896  
R:Le, S.; Sternglanz, R.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S30766  
A:Accession: S30766  
A:Molecule type: DNA  
A:Residues: 1-279 <LES>  
A:Cross-references: UNIPROT:P32447; EMBL:L07593; NID:g171090; PIDN:AAC37512.1; PID:g1710  
R:Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.  
submitted to the Protein Sequence Database, September 1995  
A:Reference number: S56891  
A:Accession: S56896  
A:Molecule type: DNA  
A:Residues: 1-279 <CZI>  
A:Cross-references: EMBL:Z49390; NID:g1008303; PIDN:CAA89410.1; PID:g1008304; MIPS:YJL11  
C:Genetics:  
A:Gene: SGD:ASF1  
A:Cross-references: SGD:S0003651; MIPS:YJL115W  
A:Map position: 10L

Query Match 84.1%; Score 69; DB 2; Length 279;  
Best Local Similarity 73.3%; Pred. No. 0.91;  
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDD 15  
:|||||:|||||  
Db 178 EEEEDDEDEDEDD 192

Search completed: September 20, 2005, 12:43:19  
Job time : 2.68911 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 15.4663 Seconds  
(without alignments)  
650.171 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEDEDEDEDDVSEGVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	26	4	AAB82810 Human low
2	136	100.0	217	2	AAW49041 Human low
3	136	100.0	217	4	AAB82803 Human low
4	136	100.0	241	3	AAB58957 Breast an
5	136	100.0	538	4	AAB82806 Human low
6	104.5	76.8	28	4	AAB82816 Rabbit lo
7	104.5	76.8	232	4	AAB82799 Rabbit lo
8	104.5	76.8	252	4	AAB82800 Rabbit lo
9	104.5	76.8	317	2	AAW49038 Rabbit lo
10	104.5	76.8	317	4	AAB82798 Rabbit lo
11	104.5	76.8	550	4	AAB82807 Rabbit lo
12	88	64.7	162	2	AAR30641 pCENP-B-1
13	88	64.7	183	3	AAB43785 Human can
14	88	64.7	183	4	AAG74067 Human col
15	88	64.7	487	8	ABM82201 Tumour-as
16	88	64.7	594	2	AAR34936 CENP-B.7
17	88	64.7	594	8	ADJ26909 Human cen
18	88	64.7	599	7	ADP65135 Human cen
19	88	64.7	599	7	ADP65197 Human cen
20	83.5	61.4	155	3	AGL17704 Arabidops
21	83.5	61.4	305	3	AAG26185 Arabidops
22	83.5	61.4	305	4	AAB67814 Amino aci
23	83.5	61.4	306	5	ABM93572 Herbicida
24	83	61.0	152	3	AAB07863 Amino aci
25	83	61.0	427	5	ABP73549 Candida a

26	83	61.0	539	3	AAV53050
27	83	61.0	1501	6	AAE36115
28	83	61.0	1953	7	ADJ80153
29	82	60.3	15	4	AAB82809
30	82	60.3	343	4	ABBS8472
31	82	60.3	358	6	ABBS84611
32	82	60.3	712	7	ADBE1623
33	82	60.3	712	7	ADDA6090
34	82	60.3	712	7	ADE57828
35	82	60.3	712	7	ADE57830
36	82	60.3	759	4	ABBS63935
37	81	59.6	444	6	ABR41713
38	81	59.6	768	8	ADP98855
39	81	59.6	1161	8	ADM57329
40	81	59.6	1162	3	AAV96255
41	81	59.6	1162	3	AAV58500
42	81	59.6	1162	4	AAB62331
43	81	59.6	1162	5	ABBS5621
44	81	59.6	1162	8	ADJ65096
45	80	58.8	77	4	AAM15201

ALIGNMENTS

RESULT 1  
AAB82810  
ID AAB82810 standard; peptide; 26 AA.  
XX AAB82810;  
XX  
DT 12-NOV-2001 (first entry)  
XX  
DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.  
XX  
KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO200164874-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US006356.  
XX  
PR 02-MAR-2000; 2000US-00517849.  
PR 14-JUL-2000; 2000US-00616289.  
(BOST-) BOSTON HEART FOUND INC.  
Lees AM, Lees RS, Law SW, Arjona AA;  
WPI; 2001-565505/63.  
New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.  
Claim 14(a); Page; 143pp; English.

The present sequence is that of a peptide comprising amino acid residues 329-354 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active fragments and analogues of LBP, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine

CC compositions. Note: the present sequence is not shown in the  
CC specification but is derived from the human LBP-2 sequence given in  
CC figure 7A (see AAB82806)  
XX  
SQ Sequence 26 AA;

Query Match 100.0%; Score 136; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 2.8e-10;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDEDDVSGSEVPESD 26  
|||||  
DB 1 EREDDDEDEDEDDVSGSEVPESD 26  
|||||

RESULT 2  
AAW49041  
ID AAW49041 standard; protein; 217 AA.  
XX  
AC AAW49041;  
XX  
DT 09-NOV-1998 (first entry)  
XX  
DE Human low density lipoprotein binding protein LBP-2.  
XX  
XW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 8..33  
FT /note= "Claim 2"  
FT Peptide 8..22  
FT /note= "Claim 2"  
FT Peptide 23..33  
FT /note= "Claim 2"  
FT Peptide 208..217  
FT /note= "Claim 2"  
XX  
PN WO9823282-A1.  
XX  
XX  
PD 04-JUN-1998.  
XX  
PF 26-NOV-1997; 97WO-US021857.  
XX  
XX 27-NOV-1996; 96US-0031930P.  
PR 03-JUN-1997; 97US-0048547P.  
XX  
PA (BOST-) BOSTON HEART FOUND INC.  
XX  
XX Lees AM, Lees RS, Law SW, Arjona AA;  
XX WPI; 1998-322455/28.  
DR N-PSDB; AAV32838.  
XX  
XX Nucleic acid encoding low density lipoprotein binding proteins and  
PT related vectors - transformed cells, proteins, and modulators of binding,  
PT useful for treatment and diagnosis of atherosclerosis and for identifying  
PT subjects at risk.  
XX  
XX Claim 1; Fig 7; 47pp; English.  
XX  
XX This polypeptide comprises novel human low density lipoprotein (LDL)  
CC binding protein LBP-2 that is capable of binding both native and methyl  
CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see  
CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human  
CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP  
CC metabolism or structure is diagnostic of a risk for atherosclerosis. The  
CC invention provides: methods for determining if an animal is at risk for  
CC atherosclerosis (e.g. for prenatal screening); methods for treating  
CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to  
CC bind LDL and thereby prevent formation of atherosclerotic plaque; and

CC methods for treating a cell having an abnormality in LBP structure or  
CC metabolism. Pharmaceutical and vaccine compositions are also provided, as  
CC well as recombinant vectors and host cells used to produce recombinant  
CC LBP  
XX  
SQ Sequence 217 AA;

Query Match 100.0%; Score 136; DB 2; Length 217;  
Best Local Similarity 100.0%; Pred. No. 2.8e-09;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDEDDVSGSEVPESD 26  
|||||  
DB 8 EREDDDEDEDEDDVSGSEVPESD 33  
|||||

RESULT 3  
AAB82803  
ID AAB82803 standard; protein; 217 AA.  
XX  
AC AAB82803;  
XX  
DT 12-NOV-2001 (first entry)  
XX  
DE Human low density lipoprotein binding protein 2 (LBP-2).  
XX  
XW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO200164874-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US006356.  
XX  
PR 02-MAR-2000; 2000US-00517849.  
PR 14-JUL-2000; 2000US-00616289.  
XX  
XX (BOST-) BOSTON HEART FOUND INC.  
XX  
XX Lees AM, Lees RS, Law SW, Arjona AA;  
XX WPI; 2001-565505/63.  
DR N-PSDB; AAB26494.  
XX  
XX New isolated low density lipoprotein binding polypeptide for treating,  
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
XX  
XX Claim 13(g); Fig 7B; 143pp; English.  
XX  
XX The present sequence is that of the N-terminal portion of novel human low  
CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
CC deduced from an isolated partial cDNA clone (see AAB26494). A full-length  
CC sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP  
CC polypeptides of the invention that are capable of binding to native and  
CC methylated low density lipoproteins. Also claimed are biologically active  
CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
CC as expression vectors, cells and methods of producing the LBPs. Methods  
CC of determining if an animal is at risk for atherosclerosis, methods for  
CC evaluating an agent for use in treating atherosclerosis, and methods for  
CC treating a cell having an abnormality in structure or metabolism of LBP  
CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
CC nucleic acid, and vaccine compositions, are also claimed  
XX  
SQ Sequence 217 AA;

Query Match 100.0%; Score 136; DB 4; Length 217;  
Best Local Similarity 100.0%; Pred. No. 2.8e-09;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDEDDVSGSEVPESD 26

```

Db      8 EEEEDDDDEDEDDVSGSEVPESD 33
      |||||||||||||||||||||||||||
RESULT 4
AAB58957
ID AAB58957 standard; protein; 241 AA.
XX
AC AAB58957;
XX
DT 27-MAR-2001 (first entry)
XX
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX
OS Homo sapiens.
XX
XX WO200055173-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005881.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-611515/58.
XX
XX N-PSDB; AAF21860.
XX
XX New human breast and ovarian cancer associated gene sequences and the
XX polypeptides encoded by these genes, useful in the prevention, treatment
XX and diagnosis of cancer, immune disorders, cardiovascular disorders and
XX neurological diseases.
XX
XX Claim 11; Page 1112; 1299pp; English.
XX
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX associated with breast and ovarian cancer. Included in the invention are
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX isolation and characterisation of the DNA and protein sequences of the
XX or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic;
XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
XX antiinflammatory; antitumor; vulnery; anticonvulsant; antibacterial;
XX antifungal; antiparasitic and cardiant activity. The polynucleotide and
XX protein sequences are used in the diagnosis of cancer, particularly
XX breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
XX and agonists may also be used in the diagnosis, prevention and treatment
XX of immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX cardiovascular disorders such as myocardial ischaemias; wound healing;
XX neurological diseases
XX
XX Sequence 241 AA;
XX
XX Query Match 100.0%; Score 136; DB 3; Length 241;
XX Best Local Similarity 100.0%; Pred. No. 3.1e-09;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Db      1 EEEEDDDDEDEDDVSGSEVPESD 26
      |||||||||||||||||||||||||||
RESULT 5
AAB82806
ID AAB82806 standard; protein; 538 AA.
XX
AC AAB82806;
XX
DT 12-NOV-2001 (first entry)
XX
DE Human low density lipoprotein binding protein 2 (LBP-2).
XX
KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX
XX Homo sapiens.
XX
XX WO200164874-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US006356.
XX
XX 02-MAR-2000; 2000US-00517849.
XX
XX 14-JUL-2000; 2000US-00616289.
XX
XX (BOST-) BOSTON HEART FOUND INC.
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX
XX WPI; 2001-565505/63.
XX
XX N-PSDB; AAB26499.
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
XX diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
XX Claim 13(j); Fig 7A; 143pp; English.
XX
XX The present sequence is that of novel human low density lipoprotein
XX binding protein 2 (LBP-2). The amino acid sequence was deduced from the
XX coding region of isolated genomic DNA (see AAB26499). It differs from the
XX sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the
XX presence of an additional 321 amino acids at the N-terminus (the cDNA
XX clone is 5' truncated). Human LBP-2 is an example of claimed LBP
XX polypeptides of the invention that are capable of binding to native and
XX methylated low density lipoproteins. Also claimed are biologically active
XX fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
XX as expression vectors, cells and methods of producing the LBPs.
XX Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529
XX -538 (see AAB82803-12) of the present sequence are claimed. Methods of
XX determining if an animal is at risk for atherosclerosis, methods for
XX evaluating an agent for use in treating atherosclerosis, and methods for
XX treating a cell having an abnormality in structure or metabolism of LBP
XX are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
XX nucleic acid, and vaccine compositions, are also claimed
XX
XX Sequence 538 AA;
XX
XX Query Match 100.0%; Score 136; DB 4; Length 538;
XX Best Local Similarity 100.0%; Pred. No. 7.5e-09;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Db      1 EEEEDDDDEDEDDVSGSEVPESD 26
      |||||||||||||||||||||||||||
RESULT 6
AAB82816
ID AAB82816 standard; peptide; 28 AA.

```





PF	26-NOV-1997;	97WO-US021857.
XX		
XX	27-NOV-1996;	96US-0031930P.
PR	03-JUN-1997;	97US-0048547P.
XX	(BOST--)	BOSTON HEART FOUND INC.
PA		
XX		
XX	Lees AM,	Lees RS, Law SW, Arjona AA;
XX		
XX	WPI; 1998-322455/28.	
DR	N-PSDB; AAV32835.	
XX		
PT	Nucleic acid encoding low density lipoprotein binding proteins and related vectors - transformed cells, proteins, and modulators of binding, useful for treatment and diagnosis of atherosclerosis and for identifying subjects at risk.	
PT		
XX		
XX	Claim 1; Fig 2; 47pp; English.	
PS		
XX		
CC	This polypeptide comprises novel rabbit low density lipoprotein (LDL) binding protein LBP-2 that is capable of binding both native and methyl LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA (see AAV22835). cDNA clones (see AAV32834-39) and encoded rabbit and human LBPs (see AA49037-42) are claimed. An abnormality in an aspect of LBP metabolism or structure is diagnostic of a risk for atherosclerosis. The invention provides methods for determining if an animal is at risk for atherosclerosis (e.g. for prenatal screening); methods for treating atherosclerosis (including gene therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent formation of atherosclerotic plaque; and methods for treating a cell having an abnormality in LBP structure or metabolism. Pharmaceutical and vaccine compositions are also provided, as well as recombinant vectors and host cells used to produce recombinant LBP	
XX		
XX	Sequence 317 AA;	
SQ		
<hr/>		
Query Match	76.8%;	Score 104.5; DB 2; Length 317;
Best Local Similarity	74.1%;	Pred. No. 3.9e-05;
Matches 20; Conservative	6;	Mismatches 0; Indels 1; Gaps 1;
<hr/>		
Qy	1 EEEEDDEDEDDED-VSEGSEVPESD 26      ::  ::	
Dd	106 EEEEEEDDDDDVVSEGSEVPESD 132	
<hr/>		
RESULT 10		
AAB82798		
ID	AAB82798 standard; protein; 317 AA.	
XX		
AC	AAB82798;	
XX		
DT	12-NOV-2001 (first entry)	
XX		
DE	Rabbit low density lipoprotein binding protein 2 (LBP-2).	
XX		
KW	Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit; atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.	
KW		
XX	Oryctolagus cuniculus.	
OS		
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 10 /note= "encoded by TAG"	
FT		
XX		
PN	WO200164874-A2.	
XX		
PD	07-SEP-2001.	
XX		
XX		
PF	28-FEB-2001; 2001WO-US006356.	
XX		
XX	02-MAR-2000; 2000US-00517849.	
PR	14-JUL-2000; 2000US-00616289.	
XX		

PA	(BOST-) BOSTON HEART FOUND INC.
XX	
PI	Lees AM, Lees RS, Law SW, Arjona AA;
XX	
XX	WPI; 2001-565505/63.
DR	N-PSDB; AAH26490.
XX	
XX	New isolated low density lipoprotein binding polypeptide for treating,
FT	diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX	
PS	Claim 13(d); Fig 4; 143pp; English.
XX	
CC	The present sequence is that of a partial sequence of novel rabbit low
CC	density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is
CC	deduced from an isolated cDNA clone (see AAH26490). A full-length
CC	sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an
CC	example of claimed polypeptides of the invention, termed LBPs, that are
CC	capable of binding to native and methylated low density lipoproteins.
CC	Also claimed are biologically active fragments and analogues of LBPs,
CC	polynucleotides encoding LBPs, as well as expression vectors, cells and
CC	methods of producing the LBPs. Methods of determining if an animal is at
CC	risk for atherosclerosis, methods for evaluating an agent for use in
CC	treating atherosclerosis, and methods for treating a cell having an
CC	abnormality in structure or metabolism of LBP are also claimed, as are
CC	pharmaceutical compositions comprising an LBP polypeptide or nucleic
CC	acid, and vaccine compositions
XX	
SQ	Sequence 252 AA;
	Query Match 76.8%; Score 104.5; DB 4; Length 252;
	Best Local Similarity 74.1%; Pred. No. 3.1e-05;
	Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;
QY	1 EEEEDDDDEDED-DD-VSEGSEVPESD 26
	:::     :::
Db	41 EEEEEEDDDDDV-VSEGSEVPESD 67
RESULT 9	
ID	AAM49038 standard; protein; 317 AA.
XX	
AC	AAM49038;
XX	
DT	09-NOV-1998 (first entry)
XX	
DE	Rabbit low density lipoprotein binding protein LBP-2.
XX	
KW	Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
KW	receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.
XX	
OS	Oryctolagus cuniculus.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 10 /note= "encoded by TAG"
FT	Protein 66..317
FT	/note= "Claim 1"
FT	Protein 86..317
FT	/note= "Claim 1"
FT	Peptide 105..132
FT	/note= "Claim 2"
FT	Peptide 105..120
FT	/note= "Claim 2"
FT	Peptide 121..132
FT	/note= "Claim 2"
FT	Peptide 211..220
FT	/note= "Claim 2"
XX	
PX	WO9823282-A1.
XX	
PD	04-JUN-1998.
XX	

PF	26-NOV-1997; 97WO-US021857.
XX	
PR	27-NOV-1996; 96US-0031930P.
XX	
PR	03-JUN-1997; 97US-0048547P.
XX	
PA	(BOST-) BOSTON HEART FOUND INC.
XX	
PI	Lees AM, Lees RS, Law SW, Arjona AA;
XX	
DR	WPI; 1998-322455/28.
DR	N-PSDB; AAV2835.
XX	
PT	Nucleic acid encoding low density lipoprotein binding proteins and
PT	related vectors - transformed cells, proteins, and modulators of binding,
PT	useful for treatment and diagnosis of atherosclerosis and for identifying
PT	subjects at risk.
XX	
PS	Claim 1; Fig 2; 47pp; English.
XX	
CC	This polypeptide comprises novel rabbit low density lipoprotein (LDL)
CC	binding protein LBP-2 that is capable of binding both native and methyl
CC	LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA
CC	(see AAV2835).. cDNA clones (see AAV2834-39) and encoded rabbit and
CC	human LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of
CC	LBP metabolism or structure is diagnostic of a risk for atherosclerosis.
CC	The invention provides: methods for determining if an animal is at risk
CC	for atherosclerosis (e.g. for prenatal screening); methods for treating
CC	atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
CC	bind LDL and thereby prevent formation of atherosclerotic plaque; and
CC	methods for treating a cell having an abnormality in LBP structure or
CC	metabolism. Pharmaceutical and vaccine compositions are also provided, as
CC	well as recombinant vectors and host cells used to produce recombinant
CC	LBP
XX	
SQ	Sequence 317 AA;
	Query Match 76.8%; Score 104.5; DB 2; Length 317;
	Best Local Similarity 74.1%; Pred. No. 3.9e-05;
	Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;
QY	1 EEEEDDDDEDED-DD-VSEGSEVPESD 26
	:::     :::
Db	106 EEEEEEDDDDDV-VSEGSEVPESD 132
RESULT 10	
ID	AAB82798 standard; protein; 317 AA.
XX	
AC	AAB82798;
XX	
DT	12-NOV-2001 (first entry)
XX	
DE	Rabbit low density lipoprotein binding protein 2 (LBP-2).
XX	
KW	Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
KW	atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX	
OS	Oryctolagus cuniculus.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 10 /note= "encoded by TAG"
FT	Protein 66..317
FT	/note= "Claim 1"
FT	Protein 86..317
FT	/note= "Claim 1"
FT	Peptide 105..132
FT	/note= "Claim 2"
FT	Peptide 105..120
FT	/note= "Claim 2"
FT	Peptide 121..132
FT	/note= "Claim 2"
FT	Peptide 211..220
FT	/note= "Claim 2"
XX	
PX	WO200164874-A2.
XX	
PD	07-SEP-2001.
XX	
PF	28-FEB-2001; 2001WO-US006356.
XX	
PR	02-MAR-2000; 2000US-00517849.
PR	14-JUL-2000; 2000US-00616289.
XX	







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 2.92778 Seconds  
(without alignments)  
854.447 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEEDDEDEDDVSEGSVEPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	64.7	599	S18735	centromere protein
2	83.5	61.4	305	T52287	probable histone d
3	82	60.3	358	S60892	nucleosome assembl
4	82	60.3	606	S70358	centromere protein
5	82	60.3	678	A54514	glutamic acid-rich
6	82	60.3	712	JH0148	nucleolin - rat
7	81	59.6	694	DNCHNL	nucleolin - chicke
8	81	59.6	1161	I59311	nardilysin (EC 3.4
9	80	58.8	707	A35804	nucleolin - human
10	79	58.1	299	A29681	nucleophosmin - Af
11	79	58.1	542	T48488	hypothetical prote
12	79	58.1	636	T51893	related to Che-1 p
13	79	58.1	802	S48529	NAB3 protein - yea
14	79	58.1	1089	S48244	NMD2 protein - yea
15	79	58.1	1105	T18295	Ap-3 adaptor compl
16	79	58.1	2206	G71611	hypothetical prote
17	78	57.4	365	F71437	probable resistanc
18	78	57.4	562	S38149	SIS2 protein - yea
19	78	57.4	792	T42963	hypothetical prote
20	77.5	57.0	1102	T84545	probable ubiquitin
21	77	56.6	896	T24169	hypothetical prote
22	77	56.6	906	T24166	hypothetical prote
23	76.5	56.2	1085	S55352	IFH1 protein - yea
24	76	55.9	677	S15667	transcription fact
25	76	55.9	864	T30441	probable capsid-as
26	76	55.9	1948	S00485	Gene 11-1 protein
27	75.5	55.5	727	S18193	autoantigen NOR-90
28	75.5	55.5	764	S09318	transcription fact
29	75	55.1	279	S30766	ASF1 protein - yea

30	75	55.1	280	2	A61047	ectodermal (ect) -
31	75	55.1	299	2	T02515	cytoskeletal prote
32	75	55.1	1132	2	T43483	translation initia
33	74.5	54.8	727	2	JC5113	ribosomal transcri
34	74.5	54.8	764	2	JC5112	ribosomal transcri
35	74.5	54.8	1133	2	T12529	hypothetical prote
36	74	54.4	121	2	T12467	hypothetical prote
37	74	54.4	147	1	RDBYUC	ubiquinol-cytochro
38	74	54.4	288	2	T22846	hypothetical prote
39	74	54.4	527	2	T18232	conserved hypochet
40	74	54.4	757	1	T50959	probable endopolyp
41	74	54.4	764	2	S64951	hypothetical prote
42	74	54.4	798	2	T33022	hypothetical prote
43	74	54.4	2094	2	S33124	tpr protein - huma
44	73.5	54.0	458	2	D88950	protein R09B5.1 [i
45	73	53.7	215	1	S01947	nonhistone chromos

ALIGNMENTS

RESULT 1

S18735

centromere protein B - human

N;Alternate names: major centromere autoantigen CENP-B

C;Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C;Accession: S18735; A27272

R;Sullivan, K.F.; Glass, C.A.

Chromosome 100, 360-370, 1991

A;Title: CENP-B is a highly conserved mammalian centromere protein with homology to the

A;Reference number: S18735; MUID:91372020; PMID:1893793

A;Accession: S18735

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-599 <SUL>

A;Cross-references: UNIPROT:P07199; EMBL:X5039; NID:g29860; PIDN:CAA38879.1; PID:g29861

R;Earnshaw, W.C.; Sullivan, K.F.; Machlin, P.S.; Cooke, C.A.; Kaiser, D.A.; Pollard, T.D

J. Cell Biol. 104, 817-829, 1987

A;Title: Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen.

A;Reference number: A27272; MUID:87166180; PMID:2435739

A;Accession: A27272

A;Molecule type: mRNA

A;Residues: 6-582, 'M', 584-591, 'LL', 594-599 <EAR>

A;Cross-references: GB:X05299; NID:g29862; PIDN:CAA28918.1; PID:g29863

C;Genetics:

A;Gene: GDB:CENPB

A;Cross-references: GDB:118768; OMIM:117140

A;Map position: 20p13-20p13

C;Keywords: DNA binding

Query Match 64.7%; Score 88; DB 2; Length 599;

Best Local Similarity 60.9%; Pred. No. 0.054;

Matches 14; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSEGSVEP 23

:::|||||:::|::|

Db 518 DDEEDEDDEDDDEEDGDEVP 540

RESULT 2

T52287

probable histone deacetylase (EC 3.5.1.-) [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004

C;Accession: T52287

R;Dangl, M.; Haas, H.; Loidl, P.

submitted to the EMBL Data Library, January 1998

A;Description: Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2.

A;Reference number: Z26016

A;Accession: T52287

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-305 <DAN>  
A;Cross-references: UNIPROT:O49209; EMBL:AF044914; PIDN:AAC02539.1  
A;Experimental source: cultivar Columbia  
C;Genetics:  
A;Gene: HD2  
C;Keywords: hydrolase

Query Match 61.4%; Score 83.5; DB 2; Length 305;  
Best Local Similarity 66.7%; Pred. No. 0.081;  
Matches 18; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

```
Qy 1 EEEEDDDEDEDEDDVSE-GSEVPESD 26
    ||: ||||| ||||| ||||| : |||
Db 156 EEDESDDDESEEDDDSEKGMVDDED 182
```

RESULT 3  
S60892  
nucleosome assembly protein 1 - soybean  
C:Species: Glycine max (soybean)  
C:Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S60892  
R:Xoon, H.W.; Kim, M.C.; Lee, S.Y.; Hwang, I.; Bahk, J.D.; Hong, J.C.; Iehimi, Y.; Cho, M.I. Gen. Genet. 249, 465-473, 1995  
A:Title: Molecular cloning and functional characterization of a cDNA encoding nucleosome  
A:Reference number: S60892; MUID:96133687; PMID:8544812  
A:Accession: S60892  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-358 <Y00>  
A:Cross-references: UNIPROT:Q39892; EMBL:L38856; NID:q1161251; PIDN:AAA88792.1; PID:q116

Query Match 60.3%; Score 82; DB 2; Length 358;  
Best Local Similarity 71.4%; Pred. No. 0.13;  
Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy	1	EEEEDEDEDEDEDDVSEGS	21
	:	:	
Dd	307	DEEDEDDEDEDEDEDEDEDE	327

RESULT 4  
S70358  
centromere protein B - Chinese hamster  
C;Species: Cricetulus griseus (Chinese hamster)  
C;Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
C;Accession: S70358  
R;Bejarano, L.A.; Valdivia, M.M.  
Biochim. Biophys. Acta 1307, 21-25, 1996  
A;Title: Molecular cloning of an intronless gene for the hamster centromere antigen CENP  
A;Reference number: S70358; MUID:96254058; PMID:8652663  
A;Accession: S70358  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-606 <BEJ>  
A;Cross-references: UNIPROT:P48988; EMBL:U20951; NID:g836955; PID:g8369  
C;Keywords: DNA binding

Query Match 60.3%; Score 82; DB 2; Length 606;  
Best Local Similarity 56.5%; Pred. No. 0.22;  
Matches 13; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 EEEEDDDDEDEEDDVSEGSEVP 23
        |:|:|:|:|:|:|:|:|:|:|
Db     525 EDEEEDEEEDDDDDDEDGDVEVP 547

```

RESULT 5  
A54514  
glutamic acid-rich protein precursor - malaria parasite (Plasmodium falciparum)  
N/Alternate names: GARP  
C/Species: Plasmodium falciparum  
C/Date: 15-Oct-1994 #sequence revision 15-Oct-1994 #text change 09-Jul-2004

C;Accession: A54514  
R;Triglia, T.; Stahl, H.D.; Crowther, P.E.; Silva, A.; Anders, R.F.; Kemp, D.J.  
Mol. Biochem. Parasitol. 31, 199-202, 1988  
A;Title: Structure of a Plasmodium falciparum gene that encodes a glutamic acid-rich prot  
A;Reference number: A54514; MUID:89040048; PMID:2903445

Query Match 60.3%; Score 82; DB 2; Length 678;  
Best Local Similarity 57.7%; Pred. No. 0.25;  
Matches 15; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

Qy      1 EEEEDDEDEDEDDVSEGSEVPESD 26
         |||::|:|||||:| | ||:
Db     640 EEDDEDEDEDEEEEEEEEEEESE 665

```

RESULT 6  
JH0148  
nucleolin - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1991 #sequence.Revision 31-Dec-1991 #text\_change 23-Jul-1999  
C:Accession: JH0148; A24088; I63110  
R:Bourbon, H.M.; Amalric, F.  
Gene 88, 187-196, 1990  
A:Title: Nucleolin gene organization in rodents: highly conserved sequences within three  
A:Reference number: JH0148; MUID:90269607; PMID:2347493

Query Match 60.3%; Score 82; DB 2; Length 712;  
Best Local Similarity 57.7%; Pred. No. 0.26;  
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

QY      1 EEEEDDEDEDEDDVSEGSEVPESD 26
        |||||:||||:| | | : |
Db     241 EEEEDDEDEDEDEDEDEDEDED 266

```

RESULT 7  
 DNCHNL  
 nucleolin - chicken  
 N;Alternate names: nucleolar protein C23  
 C;Species: Gallus gallus (chicken)  
 C;Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text\_change 09-Jul-2004  
 A;Accession: S08414; S10766; A32725; I50397; B30099  
 R;Maridor, G.; Nigg, E.A.  
 Nucleic Acids Res. 18, 1286, 1990  
 A;Title: cDNA sequences of chicken nucleolin/C23 and N038/B23, two major nucleolar proteins  
 A;Reference number: S08414; MUID:90206792; PMID:2320420  
 A;Accession: S08414  
 A;Status: translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-694 <MAR>  
 A;Cross-references: UNIPROT:P15771; EMBL:X17199; NID:G63710; PTDN:CAA35060.1; PID:G63711  
 R;Maridor, G.; Krek, W.; Nigg, E.A.  
 Biochim. Biophys. Acta 1049, 126-133, 1990  
 A;Title: Structure and developmental expression of chicken nucleolin and N038: coordinated  
 A;Reference number: S10766; MUID:90304215; PMID:2114180  
 A;Accession: S10766  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-694 <MAR2>  
 R;Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.  
 Cell 60, 791-801, 1990  
 A;Title: Identification of major nucleolar proteins as candidate mitotic substrates of cdc2  
 A;Reference number: A32725; MUID:90182668; PMID:2178776  
 A;Accession: A32725  
 A;Molecule type: protein  
 A;Residues: 56-62;63-109;197-214 <PET>  
 R;Borer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigg, E.A.  
 Cell 56, 379-390, 1989  
 A;Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.  
 A;Reference number: I50397; MUID:89119560; PMID:2914325  
 A;Accession: I50397  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 407-418, 'R', 420-519, 'T', 521-694 <BOR>  
 A;Cross-references: GB:M21791; NID:G212411; PTDN:AAA48983.1; PID:G212412  
 C;Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorganization of nucleolin; ribonucleoprotein repeat homology  
 C;Superfamily: nucleolin; ribonucleoprotein repeat homology  
 C;Keywords: DNA binding; duplication; nucleolus; nucleus; phosphoprotein; RNA binding; RNA binding; RNA binding  
 F;1-247/Domain: acidic <ACT>  
 F;54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])  
 F;254-262/Region: nuclear location signal  
 F;267-275/Region: nuclear location signal  
 F;282-347/Domain: ribonucleoprotein repeat homology <RRM1>  
 F;283-361/Domain: RNA binding #status predicted <RNA1>  
 F;372-435/Domain: ribonucleoprotein repeat homology <RRM2>  
 F;373-448/Domain: RNA binding #status predicted <RNA2>  
 F;462-525/Domain: ribonucleoprotein repeat homology <RRM3>  
 F;463-538/Domain: RNA binding #status predicted <RNA3>  
 F;554-618/Domain: ribonucleoprotein repeat homology <RRM4>  
 F;555-631/Domain: RNA binding #status predicted <RNA4>  
 F;632-694/Domain: glycine/arginine-rich <GRR>  
 F;56,63,70,77,85/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predicted

Query Match 59.6%; Score 81; DB 1; Length 694;  
 Best Local Similarity 83.3%; Pred. No. 0.32;  
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSE 18  
 |||||:|||||:|||||  
 Db 173 EEEEDDEDEDEDEDE 190

RESULT 8  
 I59311  
 N;Alternate names: N-arginine dibasic (NRD) precursor [validated] - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 A;Accession: I59311  
 R;Pierotti, A.R.; Prat, A.; Chesneau, V.; Gaudoux, F.; Leseney, A.M.; Foulon, T.; Cohen, Proc. Natl. Acad. Sci. U.S.A. 91, 6078-6082, 1994  
 A;Title: N-arginine dibasic convertase, a metalloendopeptidase as a prototype of a class of metalloendopeptidases  
 A;Reference number: I59311; MUID:94286580; PMID:8016118  
 A;Accession: I59311  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA; protein  
 A;Residues: 1-1161 <PIB>  
 A;Cross-references: UNIPROT:P47245; GB:I27124; NID:G529591; PTDN:AAA21818.1; PID:G529592  
 A;Note: the source is designated as Rattus sp. in GenBank entry RATNRDC, release 116.0  
 C;Genetics:  
 A;Gene: NRDC  
 C;Function:  
 A;Description: catalyzes the hydrolysis of peptides on the amino side of arginine-arginine  
 A;Pathway: protein degradation  
 C;Superfamily: insulin-degrading enzyme (IDE)  
 C;Keywords: hydrolase; metalloproteinase; protein degradation; zinc  
 F;244,248/Binding site: zinc (His) #status predicted  
 F;247/Active site: Glu #status predicted

Query Match 59.6%; Score 81; DB 2; Length 1161;  
 Best Local Similarity 46.2%; Pred. No. 0.53;  
 Matches 12; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGEVPE 26  
 |||||:|||||:|||||:|||||  
 Db 152 EEEEDDEDEDEDDVSEGEVPE 177

RESULT 9  
 A35804  
 nucleolin - human  
 N;Alternate names: phosphoprotein ppl00; protein B50; protein C23  
 C;Species: Homo sapiens (man)  
 C;Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text\_change 09-Jul-2004  
 A;Accession: A35804; S04631; A48138; A55996  
 R;Srivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.  
 J. Biol. Chem. 265, 14922-14931, 1990  
 A;Title: Genomic organization and chromosomal localization of the human nucleolin gene.  
 A;Reference number: A35804; MUID:90368666; PMID:2394707  
 A;Accession: A35804  
 A;Molecule type: DNA  
 A;Residues: 1-707 <SRI>  
 A;Cross-references: UNIPROT:P19338; GB:M60858; GB:J05584; NID:G189305; PTDN:AAA59954.1; PID:J05584  
 R;Srivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.  
 FEBS Lett. 250, 99-105, 1989  
 A;Title: Cloning and sequencing of the human nucleolin cDNA.  
 A;Reference number: S04631; MUID:89290043; PMID:2737305  
 A;Accession: S04631  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-707 <SR2>  
 R;Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.  
 Mol. Cell. Biol. 13, 4301-4310, 1993  
 A;Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and t(ttc)  
 A;Reference number: A48138; MUID:93309464; PMID:8321232  
 A;Accession: A48138  
 A;Molecule type: protein  
 A;Residues: 458-474 <ISH>  
 A;Experimental source: HeLa cell nuclei  
 A;Note: sequence extracted from NCBI backbone (NCBIP:134645)  
 R;Jordan, P.; Heid, H.; Kinzel, V.; Kuebler, D.  
 Biochemistry 33, 14696-14706, 1994  
 A;Title: Major cell surface-located protein substrates of an ecto-protein kinase are hom  
 A;Reference number: A55996; MUID:95086063; PMID:7993898  
 A;Accession: A55996  
 A;Molecule type: protein  
 A;Residues: 231-236;349-362;399-403;458-461;655-656, 'X', 658-660 <JOR>  
 A;Experimental source: surface-labelled HeLa cells  
 C;Genetics:

A;Gene: GDB:NCL  
A;Cross-references: GDB:125908; OMIM:164035  
A;Map position: 2q12-2qter  
C;Superfamily: nucleolin; ribonucleoprotein repeat homology  
C;Keywords: DNA binding; nucleolus; nucleus; phosphoprotein; RNA binding  
F;308-373/Domain: ribonucleoprotein repeat homology <RRM1>  
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>  
F;487-550/Domain: ribonucleoprotein repeat homology <RRM3>  
F;573-634/Domain: ribonucleoprotein repeat homology <RRM4>  
  
Query Match 58.8%; Score 80; DB 2; Length 707;  
Best Local Similarity 53.8%; Pred. No. 0.41;  
Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 EEEEDDEDEDDVSGSEVPESD 26  
:|||||:|||||:|||||:|||||:  
Db 242 DEDEDDDDDDDDDDDDDDDEEEEEE 267  
  
RESULT 10  
A29681  
nucleophosmin - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
C;Accession: A29681  
R;Schmidt-Zachmann, M.S.; Huegle-Doerr, B.; Franke, W.W.  
EMBO J. 6, 1881-1890, 1987  
A;Title: A constitutive nucleolar protein identified as a member of the nucleoplasmin family  
A;Reference number: A29681; MUID:88004394; PMID:3308448  
A;Accession: A29681  
A;Molecule type: mRNA  
A;Residues: 1-299 <SCH>  
A;Cross-references: UNIPROT:P07222; GB:X05496; NID:g64924; PIDN:CAA29046.1; PID:g64925  
C;Superfamily: nucleophosmin  
C;Keywords: molecular chaperone; nucleus  
  
Query Match 58.1%; Score 79; DB 2; Length 299;  
Best Local Similarity 56.5%; Pred. No. 0.23;  
Matches 13; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 EEEEDDEDEDDVSGSEVP 23  
:|||||:|||||:|||||:|||||:  
Db 167 DEDDDDDDDDDDDDDEEEETP 189  
  
RESULT 11  
T48488  
hypothetical protein T28J14.110 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C;Accession: T48488  
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le  
submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z24493  
A;Accession: T48488  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-542 <BEV>  
A;Cross-references: UNIPROT:Q9LYP8; EMBL:AL163652  
A;Experimental source: cultivar Columbia; BAC clone T28J14  
C;Genetics:  
A;Map position: 5  
A;Introns: 162/3; 227/1; 251/3; 299/2; 355/3; 377/2; 397/3; 419/3; 435/3; 449/2; 487/3;  
A;Note: T28J14.110  
  
Query Match 58.1%; Score 79; DB 2; Length 542;  
Best Local Similarity 46.2%; Pred. No. 0.41;  
Matches 12; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 EEEEDDEDEDDVSGSEVPESD 26  
:|||||:|||||:|||||:|||||:  
Db 129 DDDDDDDDDDDDDDDDDSKDSEVEEEE 154

RESULT 12  
T51893  
related to Che-1 protein [imported] - Neurospora crassa  
N;Alternate names: protein B23111.50  
C;Species: Neurospora crassa  
C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C;Accession: T51893  
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, August 2000  
A;Reference number: Z25858  
A;Accession: T51893  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-636 <SCH>  
A;Cross-references: EMBL:AL391572; GSPDB:GN00116; NCSP:B23111.50  
A;Experimental source: BAC clone B23111; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B23111.50  
A;Map position: 6 59/2  
A;Introns: 24/1; 59/2  
  
Query Match 58.1%; Score 79; DB 2; Length 636;  
Best Local Similarity 50.0%; Pred. No. 0.47;  
Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
  
QY 1 EEEEDDEDEDDVSGSEVPESD 26  
:|||||:|||||:|||||:|||||:  
Db 95 DEEDDEDDDDDDDDLEDGESETGSE 120  
  
RESULT 13  
S48529  
NAB3 protein - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: probable RNA/ssDNA-binding protein HMD1; protein P1945; protein YPL190c  
C;Species: Saccharomyces cerevisiae  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S48529; S65209; S60122  
R;Wilson, S.M.; Oberdorfer, A.M.; Datar, K.V.; Swedlow, J.R.; Paddy, M.R.; Swanson, M.S.  
submitted to the EMBL Data Library, January 1994  
A;Description: Characterization of Nuclear Polyadenylated RNA-Binding Proteins from Sacca  
A;Reference number: S48529  
A;Accession: S48529  
A;Molecule type: DNA  
A;Residues: 1-802 <WIL>  
A;Cross-references: UNIPROT:P38996; EMBL:U05314; NID:g476219; PID:g476220  
R;Rieger, M.; Mueller-Auer, S.; Schaefer, M.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S65202  
A;Accession: S65209  
A;Molecule type: DNA  
A;Residues: 1-802 <RIE>  
A;Cross-references: EMBL:Z73546; NID:g1370396; PID:g1370397; GSPDB:GN00016; MIPS:YPL190c  
A;Experimental source: strain S288C (AB972)  
R;Sugimoto, K.; Matsumoto, K.; Kornberg, R.D.; Reed, S.I.; Wittenberg, C.  
Mol. Gen. Genet. 248, 712-718, 1995  
A;Title: Dosage suppressors of the dominant G1 cyclin mutant CLN3-2: identification of a  
A;Reference number: S60122; MUID:96069710; PMID:7476874  
A;Accession: S60122  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-340, 1', 342-802 <SUG>  
A;Cross-references: GB:D37935; NID:g1235749; PID:g1235750  
C;Genetics:  
A;Gene: SGD:NAB3; NAB3; HMD1; MIPS:YPL190c  
A;Cross-references: SGD:S0006111  
A;Map position: 16L  
C;Superfamily: yeast NAB3 protein; ribonucleoprotein repeat homology  
C;Keywords: nucleus; RNA binding  
F;331-396/Domain: ribonucleoprotein repeat homology <RRM1>  
  
Query Match 58.1%; Score 79; DB 1; Length 802;  
Best Local Similarity 53.8%; Pred. No. 0.59;



0;

04

ed mRNA decay p

---

ral -1 frameship

---

©

0

04  
Zhen, L.; Rusi  
ltered in the m

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 13.2069 Seconds  
(without alignments)  
1008.117 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEEDDEDEDDVSEGSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	100.0	236	2 Q961M4	Q961m4 homo sapien
2	136	100.0	285	2 Q6P0R3	Q6p0r3 homo sapien
3	136	100.0	295	2 Q6P1S7	Q6p1s7 homo sapien
4	136	100.0	538	2 Q6SPF0	Q6spf0 homo sapien
5	104.5	76.8	550	2 Q6SPF9	Q6spe9 cryotolagus
6	97	71.3	239	1 CENB_SHEEP	P49451 ovis aries
7	88	64.7	138	2 Q96E14	Q96e14 homo sapien
8	88	64.7	599	1 CENB_HUMAN	P07199 homo sapien
9	87	64.0	1161	1 NRDC_MOUSE	Q8bhg1 mus musculus
10	87	64.0	1161	2 Q8R320	Q8r320 mus musculus
11	84	61.8	490	2 Q81E75	Q81e75 plasmodium
12	84	61.8	976	2 Q9DUN0	Q9dun0 human herpe
13	84	61.8	1036	2 Q9DUM3	Q9dum3 human herpe
14	84	61.8	2246	2 Q8T253	Q8t253 dictyosteli
15	83.5	61.4	305	2 Q49209	Q49209 arabidopsis
16	83.5	61.4	306	2 Q9FNJ6	Q9fnj6 arabidopsis
17	83	61.0	300	2 Q7R5A5	Q7r5a5 plasmodium
18	83	61.0	374	2 Q7P1T7	Q7p1t7 anopheles g
19	83	61.0	374	2 Q70Z19	Q70z19 nicotiana t
20	83	61.0	457	2 Q6P440	Q6p440 homo sapien
21	83	61.0	492	2 Q6DKH9	Q6dkh9 homo sapien
22	83	61.0	1158	2 Q7Q7D6	Q7q7d6 anopheles g
23	83	61.0	1544	2 Q6ZNL7	Q6znl7 homo sapien
24	83	61.0	2004	1 CHD8_HUMAN	Q9hck8 homo sapien
25	83	61.0	2275	2 Q81BX2	Q81bx2 plasmodium
26	83	61.0	2302	2 Q68DQ0	Q68dq0 homo sapien
27	82	60.3	235	2 Q7RCY6	Q7rcy6 plasmodium
28	82	60.3	343	2 Q9W3B2	Q9w3b2 drosophila
29	82	60.3	358	2 Q39892	Q39892 glycine max
30	82	60.3	484	2 Q7RH75	Q7rh75 plasmodium
31	82	60.3	606	1 CENB_CRIGR	P48988 cricetus

#### RESULT 1

ID	Q961M4	PRELIMINARY;	PRT;	236 AA.
AC	Q961M4;			
DT	01-DEC-2001 (Tremblrel. 19, Created)			
DT	01-MAR-2004 (Tremblrel. 26, Last sequence update)			
DT	25-OCT-2004 (Tremblrel. 28, Last annotation update)			
DE	LOC90378 protein (Hypothetical protein) (Fragment).			
GN	Name=LOC90378;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymph;			
RX	MEDLINE=2238857; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymph;			
RA	Strausberg R.;			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymph;			
RA	Director MGC Project;			
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC007384; AAH07384.2; -			
DR	EMBL; BC080588; AAH080588.1; -			
DR	HSSP; P39769; 1KW4.			
DR	InterPro; IPR001660; SAM.			
DR	InterPro; IPR010993; SAM_homology.			
DR	Pfam; PF00536; SAM_1; 1.			
DR	SMART; SM00454; SAM; 1.			

#### ALIGNMENTS

32	82	60.3	678	1	GARP_PLAFF
33	82	60.3	712	1	NUCL_RAT
34	82	60.3	759	2	Q9VX49
35	82	60.3	1364	2	Q8T245
36	81.5	59.9	599	1	CENB_MOUSE
37	81.5	59.9	599	2	Q7TSG8
38	81.5	59.9	622	2	Q923C5
39	81	59.6	169	2	Q6KH7
40	81	59.6	215	2	Q9PUK9
41	81	59.6	214	2	Q9VH06
42	81	59.6	320	2	Q61Q75
43	81	59.6	320	2	Q7ZU28
44	81	59.6	416	2	Q9YPA9
45	81	59.6	694	1	NUCL_CHICK

P13816 plasmodium  
P13383 rattus norv  
Q9VX49 drosophila  
Q8T245 dictyosteli  
P27790 mus musculus  
Q7T898 mus musculus  
Q923C5 mus musculus  
Q6KH7 mycoplasma  
Q9PUK9 gallus gall  
Q9VH06 gallus gall  
Q61Q75 brachydanio  
Q7ZU28 brachydanio  
Q9YPA9 human herpe  
P15771 gallus gall

```

DR PROSITE: PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 136; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 27 EEEEDDEDEDDVSGSEVPESD 52
   |||||

RESULT 2
Q6P0R3 PRELIMINARY; PRT; 285 AA.
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC065477; AAH65477.1; -.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM 1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2P936CAF11F901 CRC64;

Query Match 100.0%; Score 136; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 3.8e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 76 EEEEDDEDEDDVSGSEVPESD 101
   |||||

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC030129; AAH30129.1; -.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM 1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 136; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
   |||||
Db 86 EEEEDDEDEDDVSGSEVPESD 111
   |||||

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Aorta;
RA  Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL  Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AY453840; AAR24087.1; -.
DR  InterPro: IPR001660; SAM.
DR  Pfam: PF00536; SAM_1; 1.
DR  SMART: SM00454; SAM; 1.
DR  PROSITE: PS0105; SAM_DOMAIN; 1.
SQ  SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 136; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 EEEEDDEDEDEDDVSEGSEVPESD 26
Db  329 EEEEDDEDEDEDDVSEGSEVPESD 354

RESULT 5
Q6SPE9 PRELIMINARY; PRT; 550 AA.
AC  Q6SPE9;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  Atherin.
OS  Oryctolagus cuniculus (Rabbit).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX  NCBI_TaxID=9986;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Aorta;
RA  Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL  Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AY453841; AAR24088.1; -.
DR  InterPro: IPR001660; SAM.
DR  Pfam: PF00536; SAM_1; 1.
DR  SMART: SM00454; SAM; 1.
DR  PROSITE: PS0105; SAM_DOMAIN; 1.
SQ  SEQUENCE 550 AA; 57019 MW; AF1CF2B780D879A7 CRC64;

Query Match 76.8%; Score 104.5; DB 2; Length 550;
Best Local Similarity 74.1%; Pred. No. 0.01;
Matches 20; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

QY  1 EEEEDDEDEDEDD-VSEGSEVPESD 26
Db  339 EEEEDDEDEDDVSEGSEVPESD 365

RESULT 6
CENB_SHEEP STANDARD; PRT; 239 AA.
AC  P494E1;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Major centromere autoantigen B (Centromere protein B) (CENP-B)
DE  (Fragment).
GN  Names=CENB;
OS  Ovis aries (Sheep).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Caprinae; Ovis.
OX  NCBI_TaxID=9940;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Lung;
RX  MEDLINE=97049078; PubMed=8893808;

RA  Burkin D.J., Jones C.A., Burkin H.R., McGrew J.A., Broad T.E.;
RT  "Sheep CENPB and CENPC genes show a high level of sequence similarity
RL  and conserved synteny with their human homologs.";
RL  Cytogenet. Cell Genet. 74:86-89(1996).
CC  -1- FUNCTION: Interacts with centromeric heterochromatin in
CC  chromosomes and binds to a specific subset of alphoid satellite
CC  DNA, called the CENP-B box. May organize arrays of centromere
CC  satellite DNA into a higher order structure which then directs
CC  centromere formation and kinetochore assembly in mammalian
CC  chromosomes (By similarity).
CC  -1- SUBUNIT: Homodimer (By similarity).
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- SIMILARITY: Contains 1 CENPB domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: U35655; AAA79098.1; -.
DR  InterPro: IPR004875; CENP-B.
KW  Centromere; Chromosomal protein; DNA-binding; Nuclear protein.
FT  NON_TER 1
SQ  SEQUENCE 239 AA; 26436 MW; 259C6C72E7D9C135 CRC64;

Query Match 71.3%; Score 97; DB 1; Length 239;
Best Local Similarity 78.3%; Pred. No. 0.025;
Matches 18; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 EEEEDDEDEDEDDVSEGSEVP 23
Db  158 EEEEDDEDEDDDEEDDEVP 180

RESULT 7
Q96E14 PRELIMINARY; PRT; 138 AA.
AC  Q96E14;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Hypothetical protein (fragment).
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Placenta;
RA  Strausberg R.;
RL  Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL: BC012297; AAH12297.1; -.
DR  HSSP: P07199; IUFF.
KW  Hypothetical protein.
FT  NON_TER 1
SQ  SEQUENCE 138 AA; 15273 MW; 5287985028B9D41D CRC64;

Query Match 64.7%; Score 88; DB 2; Length 138;
Best Local Similarity 60.9%; Pred. No. 0.11;
Matches 14; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY  1 EEEEDDEDEDEDDVSEGSEVP 23
Db  57 DDEEDDEDEDDDEEDGDEVP 79

RESULT 8
CENB_HUMAN STANDARD; PRT; 599 AA.
ID  CENB_HUMAN
AC  P07139;

```

DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Major centromere autoantigen B (Centromere protein B) (CENP-B).  
 GN Name=CENPB;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91372020; PubMed=1893793;  
 RA Sullivan K.F., Glass C.A.;  
 RT "CENP-B is a highly conserved mammalian centromere protein with  
 RT homology to the helix-loop-helix family of proteins.";  
 RL Chromosoma 100:360-370(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;  
 RA Deloukas P., Matthews L.H., Ashurst J.L., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Dedman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leharstaiho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.B., Martin S.L., McConachie L.J., McLeay K., McKerray A.A.,  
 RA Milne S.A., Mistry D., Moore M.C., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.M., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.D., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 6-599 FROM N.A.  
 RX MEDLINE=87166180; PubMed=2435739; DOI=10.1083/jcb.104.4.817;  
 RA

RA Earnshaw W.C., Sullivan K.F., Machlin P.S., Cooke C.A., Kaiser D.A.,  
 RA Pollard T.D., Rothfield N.F., Cleveland D.W.;  
 RT "Molecular cloning of cDNA for CENP-B, the major human centromere  
 RT autoantigen.";  
 RL J. Cell Biol. 104:817-829(1987).  
 RN [5]  
 RP SUBUNITS, AND DOMAINS.  
 RX MEDLINE=93107144; PubMed=1469042; DOI=10.1083/jcb.119.6.1413;  
 RA Yoda K., Kitagawa K., Masumoto H., Muro Y., Okazaki T.;  
 RT "A human centromere protein, CENP-B, has a DNA binding domain  
 RT containing four potential alpha helices at the NH2 terminus, which is  
 RT separable from dimerizing activity.";  
 RL J. Cell Biol. 119:1413-1427(1992).  
 RN [6]  
 RP STRUCTURE BY NMR OF 1-56.  
 RX MEDLINE=98119825; PubMed=9451007; DOI=10.1093/emboj/17.3.827;  
 RA Iwahara J., Kigawa T., Kitagawa K., Masumoto H., Okazaki T.,  
 RA Yokoyama S.;  
 RT "A helix-turn-helix structure unit in human centromere protein B  
 RT (CENP-B).";  
 RL EMBO J. 17:827-837(1998).  
 CC -!- FUNCTION: Interacts with centromeric heterochromatin in  
 CC chromosomes and binds to a specific subset of aliphoid satellite  
 CC DNA, called the CENP-B box. May organize arrays of centromere  
 CC satellite DNA into a higher order structure which then directs  
 CC centromere formation and kinetochore assembly in mammalian  
 CC chromosomes.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Contains 1 CENPB domain.  
 CC -!- SIMILARITY: Contains 1 HTH psq-type DNA-binding domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X55039; CAA38879.1; -.  
 DR EMBL; AL109804; CAC17547.1; -.  
 DR EMBL; BC053847; AAH53847.1; -.  
 DR EMBL; X05299; CAA28918.1; -.  
 DR FTR; S18735; S18735.  
 DR PDB; 1BW6; NMR; A=1-56.  
 DR PDB; 1HLV; X-ray; A=1-131.  
 DR PDB; 1UFI; X-ray; A/B/C/D=536-599.  
 DR Genew; HGNC:1852; CENPB.  
 DR MIM; 117140; -.  
 DR GO; GO:000775; C:chromosome, pericentric region; NAS.  
 DR GO; GO:0003682; F:chromatin binding; NAS.  
 DR GO; GO:0003696; F:satellite DNA binding; NAS.  
 DR GO; GO:0000069; P:centromere/kinetochore complex maturation; NAS.  
 DR InterPro; IPR004875; CENP-B.  
 DR InterPro; IPR006695; CENP-B\_N.  
 DR InterPro; IPR006600; CENPB.  
 DR InterPro; IPR009057; Homeodomain-like.  
 DR Pfam; PF04218; CENP-B\_N; 1.  
 DR Pfam; PF03184; DDE; 1.  
 DR SMART; SM00674; CENPB; 1.  
 DR PROSITE; PS50960; HTH\_PSQ; 1.  
 DR 3D-structure; Centromere; Chromosomal protein; DNA-binding;  
 DR Nuclear protein.  
 KW DOMAIN 1 52 HTH psq-type.  
 FT DNA\_BIND 28 48 H-T-H motif.  
 FT DNA\_BIND 97 129 H-T-H motif.  
 FT DOMAIN 404 465 Glu-rich (acidic).  
 FT DOMAIN 508 538 Asp/Glu-rich (acidic).  
 FT CONFLICT 583 583 R -> M (in Ref. 4).  
 FT CONFLICT 592 593 VR -> LL (in Ref. 4).  
 FT HELIX 10 22  
 FT TURN 24 25

```

FT HELIX 28 35
FT TURN 36 36
FT HELIX 37 47
FT TURN 39 47
FT TURN 48 48
FT TURN 49 59
FT HELIX 60 64
FT TURN 65 65
FT TURN 73 74
FT HELIX 75 88
FT TURN 79 91
FT HELIX 92 92
FT TURN 97 111
FT TURN 112 112
FT TURN 114 115
FT HELIX 120 129
FT TURN 130 130
SQ SEQUENCE 599 AA; 65171 MW; 9B4B7DB957A9148A CRC64;

Query Match 64.7%; Score 88; DB 1; Length 599;
Best Local Similarity 60.0%; Pred. No. 0.48;
Matches 14; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEBDDDDDEDDDDVSGSEVPSD 23
DB 518 DDEBDDDDDDDDDDSGDGEVP 540

RESULT 9
NRDC MOUSE STANDARD; PRT; 1161 AA.
AC QBHGL;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nardilysin precursor (EC 3.4.24.61) (N-arginine dibasic convertase)
DE (NRD convertase) (NRD-C).
GN Name=Nrdl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasakawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saio R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Regi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Buit C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagahima T., Numata K., Okido T., Pavan W.J., Perte G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Varazdo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."

```

```

RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klautner R.D., Collins P.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -I- FUNCTION: Cleaves peptide substrates on the N-terminus of arginine
CC residues in dibasic pairs.
CC -I- CATALYTIC ACTIVITY: Hydrolysis of polypeptides, preferably at
CC Xaa-|-Arg-Lys, and less commonly at Arg-|-Arg-Xaa, in which Xaa is
CC not Arg or Lys.
CC -I- COFACTOR: Zinc (Probable).
CC -I- SIMILARITY: Belongs to the peptidase M16 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK031548; BAC27445.1; -
DR EMBL; BC036128; AAR36128.1; -
DR MEROPS; M16.005; -
DR MGD; MGI:1201386; Nrdl.
DR InterPro; IPR001431; Insulinase like.
DR InterPro; IPR007883; Peptidase_M16_C.
DR Pfam; PF00675; Peptidase M16; 1.
DR Pfam; PF05193; Peptidase M16 C; 2.
DR PROSITE; PS00143; INSULINASE; 1.
KW Hydrolase; Metalloprotease; Signal; Zinc.
FT SIGNAL 1 18 Potential.
FT CHAIN 19 1161 Nardilysin.
FT DOMAIN 139 209 Asp/Glu-rich (highly acidic).
FT DOMAIN 140 151 Poly-Glu.
FT DOMAIN 152 165 Poly-Asp.
FT DOMAIN 185 190 Poly-Asp.
FT DOMAIN 193 196 Poly-Asp.
FT METAL 244 244 Zinc (By similarity).
FT ACT_SITE 247 247 By similarity.
FT METAL 248 248 Zinc (By similarity).
FT METAL 325 325 Zinc (By similarity).
SQ SEQUENCE 1161 AA; 132890 MW; 21334221632A5122 CRC64;

Query Match 64.0%; Score 87; DB 1; Length 1161;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 13; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 BEBDDDDDEDDDDVSGSEVPSD 26
DB 149 BEBDDDDDDDDDDDDSGABIQDD 174

```

```
RESULT 10
Q8R320
ID Q8R320 PRELIMINARY; PRT; 1161 AA.
AC Q8R320;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nardillysin, N-arginine dibasic convertase, NRD convertase 1.
GN Name=Nrd1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smillius D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026832; AAH26832.1; -.
DR MGD; MGI:1201386; Nrd1.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001431; Insulinase like.
DR InterPro; IPR007863; Peptidase_M16_C.
DR Pfam; PF00675; Peptidase_M16; 1.
DR Pfam; PF05193; Peptidase_M16_C; 2.
DR PROSITE; PS00143; INSULINASE_1.
SQ SEQUENCE 1161 AA; 132876 MW; 4A5F3221632A5073 CRC64;

Query Match 64.0%; Score 87; DB 2; Length 1161;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 13; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVDVSGSEVPESD 26
|||:||||:||||:||||:|
Db 149 EEEEDDDDDDDDDDDSDSGAEIQDD 174

RESULT 11
Q8IE75
ID Q8IE75 PRELIMINARY; PRT; 490 AA.
AC Q8IE75;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Microfibril-associated protein homologue, putative.
GN Name=MAI131.132;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

us-10-671-242-20.rup
```

```
OX NCBI_TaxID=36329;
RN [1]
SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52390.1; -.
DR GO; GO:0001527; C:microfibril; IEA.
DR InterPro; IPR009730; MFAP1_C.
DR Pfam; PF06991; MFAP1_C; 1.
SQ SEQUENCE 490 AA; 59499 MW; 62CFED2EDC0C0175 CRC64;

Query Match 61.8%; Score 84; DB 2; Length 490;
Best Local Similarity 53.8%; Pred. No. 0.99;
Matches 14; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVDVSGSEVPESD 26
|||:||||:||||:||||:|
Db 205 EDEDDDDDDDDDDDDDDDDPSD 230

RESULT 12
Q9DUN0
ID Q9DUN0 PRELIMINARY; PRT; 976 AA.
AC Q9DUN0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Orf73.
OS Human herpesvirus 8 (HHV-8) (Kaposi's sarcoma-associated herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=PK-1;
RX MEDLINE=99445611; PubMed=10515805;
RA Gao S.J., Zhang Y.J., Deng J.H., Rabkin C.S., Flore O., Jenson H.B.;
RT "Molecular polymorphism of Kaposi's sarcoma-associated herpesvirus
(R human herpesvirus 8) latent nuclear antigen: evidence for a large
RT repertoire of viral genotypes and dual infection with different viral
RT genotypes.";
RL J. Infect. Dis. 180:1466-1476(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=PK-1;
RX MEDLINE=20381179; PubMed=10900044;
RA Zhang Y.J., Deng J.H., Rabkin C., Gao S.J.;
RT "Hot-spot variations of Kaposi's sarcoma-associated herpesvirus latent
RT nuclear antigen and application in genotyping by PCR-RFLP.";
RL J. Gen. Virol. 81:2049-2058(2000).
DR EMBL; AF192756; AAG01636.1; -.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYSIN.
SQ SEQUENCE 976 AA; 112017 MW; E5781E2A509FF70B CRC64;

Query Match 61.8%; Score 84; DB 2; Length 976;
Best Local Similarity 61.5%; Pred. No. 1.9;
Matches 16; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVDVSGSEVPESD 26
|||:||||:||||:||||:|
Db 382 EEEEDDEDEDDDDDDDDDEEDD 407

RESULT 13
Q9DUM3
ID Q9DUM3 PRELIMINARY; PRT; 1036 AA.
AC Q9DUM3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
```



```

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Latent nuclear antigen (Fragment).
OS Human herpesvirus 8 (HHV-8) (Kaposi's sarcoma-associated herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21165304; PubMed=11264383;
RZ DOI=10.1128/JVI.75.8.3948-3959.2001;
RA Plozet T., Tramier M., Coppey M., Nicolas J.C., Marechal V.;
RT "Close but distinct regions of human herpesvirus 8 latency-associated
RT nuclear antigen 1 are responsible for nuclear targeting and binding to
RT human mitotic chromosomes."
RL J. Virol. 75:3948-3959(2001).
DR EMBL; AF305694; AAG27458.1; -.
DR HSSP; P02649; INPN.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR000533; Tropomyosin.
DR PRINTS; PR00194; TROPOMYOSIN.
FT NON TER 1036 1036
SQ SEQUENCE 1036 AA; 119328 MW; 2959EDD2C1C042B8 CRC64;

Query Match 61.8%; Score 84; DB 2; Length 1036;
Best Local Similarity 61.5%; Pred. No. 2.1;
Matches 16; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGSVEPESD 26
Db 402 EEEEDDEDEDEDDDDDEDEDEDD 427

RESULT 14
Q87253
ID Q87253 PRELIMINARY; PRT; 2246 AA.
AC Q87253;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szatranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115683; AAL92693.2; -.
DR InterPro; IPR002110; ANK.
DR SMART; SM00248; ANK; 5.
KW Hypothetical protein.
SQ SEQUENCE 2246 AA; 265604 MW; E99C27A2E19FC67A CRC64;

Query Match 61.8%; Score 84; DB 2; Length 2246;
Best Local Similarity 57.7%; Pred. No. 4.4;
Matches 15; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSEGSVEPESD 26
Db 330 EEEEDDEDEDEDDDDDDDEDEDD 355

RESULT 15

```

```

O49209
ID O49209 PRELIMINARY; PRT; 305 AA.
AC O49209;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Putative histone deacetylase.
DE Name=HD2; Synonyms=HD2B;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Dangl M., Haas H., Loidl P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20252855; PubMed=10792817;
RA Wu K., Tian L., Malik K., Brown D., Miki B.;
RT "Functional analysis of HD2 histone deacetylase homologues in
RT Arabidopsis thaliana."
RL Plant J. 22:19-27(2000).
DR EMBL; AF044914; AAC02539.1; -.
DR EMBL; AF195546; AAG28473.1; -.
DR PIR; T52287; T52287.
DR PIR; T52287; T52287.
SQ SEQUENCE 305 AA; 32219 MW; 58AA6D5764930944 CRC64;

Query Match 61.4%; Score 83.5; DB 2; Length 305;
Best Local Similarity 66.7%; Pred. No. 0.7;
Matches 18; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 EEEEDDEDEDEDDVSE-GSEVPESD 26
Db 156 EEEEDDEDESEDDDDSEKGMVDDEDD 182

Search completed: September 20, 2005, 12:41:44
Job time : 14.2069 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 14.3525 Seconds  
(without alignments)  
733.538 Million cell updates/sec

Title: US-10-671-242-20

Perfect score: 136

Sequence: 1 EEEEDDEDEDDVSGSEVPESD 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	136	100.0	26	9	US-09-962-055-20 Sequence 20, Appl
2	136	100.0	26	9	US-09-976-740-20 Sequence 20, Appl
3	136	100.0	26	13	US-10-023-523-20 Sequence 20, Appl
4	136	100.0	26	13	US-10-023-523-20 Sequence 20, Appl
5	136	100.0	26	15	US-10-616-187-20 Sequence 20, Appl
6	136	100.0	26	15	US-10-671-242-20 Sequence 20, Appl
7	136	100.0	217	9	US-09-962-055-7 Sequence 7, Appli
8	136	100.0	217	9	US-09-976-740-7 Sequence 7, Appli
9	136	100.0	217	13	US-10-023-523-7 Sequence 7, Appli
10	136	100.0	217	13	US-10-023-523-7 Sequence 7, Appli
11	136	100.0	217	15	US-10-616-187-7 Sequence 7, Appli

12	136	100.0	217	15	US-10-671-242-7	Sequence 7, Appli
13	136	100.0	241	9	US-09-925-298-665	Sequence 665, App
14	136	100.0	241	14	US-10-102-806-665	Sequence 665, App
15	136	100.0	538	9	US-09-976-740-43	Sequence 43, Appl
16	136	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
17	136	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
18	136	100.0	538	15	US-10-616-187-43	Sequence 43, Appl
19	136	100.0	538	15	US-10-671-242-43	Sequence 43, Appl
20	104.5	76.8	28	9	US-09-962-055-26	Sequence 26, Appl
21	104.5	76.8	28	9	US-09-976-740-26	Sequence 26, Appl
22	104.5	76.8	28	13	US-10-023-523-26	Sequence 26, Appl
23	104.5	76.8	28	13	US-10-023-523-26	Sequence 26, Appl
24	104.5	76.8	28	15	US-10-616-187-26	Sequence 26, Appl
25	104.5	76.8	28	15	US-10-671-242-26	Sequence 26, Appl
26	104.5	76.8	232	9	US-09-962-055-3	Sequence 3, Appli
27	104.5	76.8	232	9	US-09-976-740-3	Sequence 3, Appli
28	104.5	76.8	232	13	US-10-023-523-3	Sequence 3, Appli
29	104.5	76.8	232	13	US-10-023-523-3	Sequence 3, Appli
30	104.5	76.8	232	15	US-10-616-187-3	Sequence 3, Appli
31	104.5	76.8	232	15	US-10-671-242-3	Sequence 3, Appli
32	104.5	76.8	252	9	US-09-962-055-4	Sequence 4, Appli
33	104.5	76.8	252	9	US-09-976-740-4	Sequence 4, Appli
34	104.5	76.8	252	13	US-10-023-523-4	Sequence 4, Appli
35	104.5	76.8	252	13	US-10-023-523-4	Sequence 4, Appli
36	104.5	76.8	252	15	US-10-616-187-4	Sequence 4, Appli
37	104.5	76.8	317	9	US-09-962-055-2	Sequence 2, Appli
38	104.5	76.8	317	9	US-09-976-740-2	Sequence 2, Appli
39	104.5	76.8	317	13	US-10-023-523-2	Sequence 2, Appli
40	104.5	76.8	317	13	US-10-023-523-2	Sequence 2, Appli
41	104.5	76.8	317	15	US-10-616-187-2	Sequence 2, Appli
42	104.5	76.8	317	15	US-10-671-242-2	Sequence 2, Appli
43	104.5	76.8	550	9	US-09-976-740-47	Sequence 47, Appl
44	104.5	76.8	550	9	US-10-023-523-47	Sequence 47, Appl
45	104.5	76.8	550	13	US-10-023-523-47	Sequence 47, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-962-055-20

; Sequence 20, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 136; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 2
US-09-976-740-20
; Sequence 20, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-20

Query Match      100.0%; Score 136; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 3
US-10-023-529-20
; Sequence 20, Application US/10023529
; Publication No. US20020129368A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 136; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26

RESULT 4
US-10-023-523-20
; Sequence 20, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-20

Query Match      100.0%; Score 136; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
Db 1 EEEEDDEDEDDVSGSEVPESD 26
```

```
RESULT 5
US-10-616-187-20
; Sequence 20, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-20

Query Match      100.0%; Score 136; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
Db 1 EREDDDEDEDDVSGSEVPESD 26

RESULT 6
US-10-671-242-20
; Sequence 20, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-20

Query Match      100.0%; Score 136; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
Db 1 EREDDDEDEDDVSGSEVPESD 26

RESULT 7
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION/DOCKET NUMBER: 35,965
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match      100.0%; Score 136; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
Db 8 EREDDDEDEDDVSGSEVPESD 33

RESULT 8
US-09-976-740-7
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-20

Query Match      100.0%; Score 136; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
Db 1 EREDDDEDEDDVSGSEVPESD 26

RESULT 7
US-09-962-055-7
; Sequence 7, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION/DOCKET NUMBER: 35,965
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-962-055-7

Query Match      100.0%; Score 136; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EREDDDEDEDDVSGSEVPESD 26
Db 8 EREDDDEDEDDVSGSEVPESD 33

RESULT 8
US-09-976-740-7
```

```
; Sequence 7, Application US/09976740
; Publication NO. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-7

Query Match      100.0%; Score 136; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
DB 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 9
US-10-023-529-7
; Sequence 7, Application US/10023529
; Publication NO. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 136; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
DB 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 10
US-10-023-523-7
; Sequence 7, Application US/10023523
; Publication NO. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-7

Query Match      100.0%; Score 136; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDDVSGSEVPESD 26
DB 8 EEEEDDEDEDDVSGSEVPESD 33

RESULT 11
US-10-616-187-7
; Sequence 7, Application US/10616187
; Publication NO. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
```

```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-7

Query Match      100.0%; Score 136; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 8 EEEEDDEDEDEDDVSGSEVPESD 33

RESULT 12
US-10-671-242-7
; Sequence 7, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 136; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 8 EEEEDDEDEDEDDVSGSEVPESD 33

RESULT 13
US-09-925-298-665
; Sequence 665, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 136; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 8 EEEEDDEDEDEDDVSGSEVPESD 33

RESULT 14
US-10-102-806-665
; Sequence 665, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 665
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-665

Query Match      100.0%; Score 136; DB 9; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26
   |||||
Db 32 EEEEDDEDEDEDDVSGSEVPESD 57

RESULT 15
US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
```

; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/976,740  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 538  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-740-43

Query Match 100.0%; Score 136; DB 9; Length 538;  
Best Local Similarity 100.0%; Pred. No. 4,1e-07;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEEEDDEDEDEDDVSGSEVPESD 26  
|||||  
Db 329 EEEEDDEDEDEDDVSGSEVPESD 354

Search completed: September 20, 2005, 12:53:06  
Job time : 14.3525 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 6.54345 Seconds  
(without alignments)  
650.171 Million cell updates/sec

Title: US-10-671-242-21

Perfect score: 54

Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	54	100.0	11	4 AAB82811	Aab82811 Human low
2	54	100.0	12	4 AAB82817	Aab82817 Rabbit lo
3	54	100.0	26	4 AAB82810	Aab82810 Human low
4	54	100.0	28	4 AAB82816	Aab82816 Rabbit lo
5	54	100.0	217	2 AAW49041	Aaw49041 Human low
6	54	100.0	217	4 AAB82803	Aab82803 Human low
7	54	100.0	232	4 AAB82799	Aab82799 Rabbit lo
8	54	100.0	241	3 AAB58957	Aab58957 Breast an
9	54	100.0	252	4 AAB82800	Aab82800 Rabbit lo
10	54	100.0	317	2 AAW49038	Aaw49038 Rabbit lo
11	54	100.0	317	4 AAB82798	Aab82798 Rabbit lo
12	54	100.0	538	4 AAB82806	Aab82806 Human low
13	54	100.0	550	4 AAB82807	Aab82807 Rabbit lo
14	38	70.4	759	7 AAB85225	Abm85225 Human pro
15	38	70.4	808	8 ADM24118	Adm24118 Bacterial
16	38	70.4	906	2 AAY27126	Aay27126 Amino aci
17	38	70.4	906	6 AAM79220	Aam79220 Human pro
18	38	70.4	906	6 ABU03746	Abu03746 Human exp
19	38	70.4	906	6 ABU03752	Abu03752 Human exp
20	38	70.4	906	6 ABU03755	Abu03755 Human exp
21	38	70.4	906	6 ABU03750	Abu03750 Human exp
22	38	70.4	906	6 ABU03754	Abu03754 Human exp
23	38	70.4	906	6 AAE38431	Aae38431 Human Cbl
24	38	70.4	906	7 ADB80950	Adb80950 RING-SH c
25	38	70.4	906	8 ADJ66658	Adj66658 c-CBL pro

ALIGNMENTS

RESULT 1

AAB82811

ID AAB82811 standard; peptide; 11 AA.

XX AAB82811;

DT 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 2 (LBP-2) peptide.

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;

KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

(BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 14(a); Page; 143pp; English.

The present sequence is that of a peptide comprising amino acid residues 344-354 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine

CC compositions. Note: the present sequence is not shown in the  
 CC specification but is derived from the human LBP-2 sequence given in  
 CC figure 7A (see AAB82806)

XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0039; Mismatches 0; Indels 0; Gaps 0;  
 Matches 11; Conservative 0;

QY 1 VSEGSSEVPESD 11  
 |||||  
 DB 1 VSEGSSEVPESD 11

RESULT 2  
 AAB82817  
 ID AAB82817 standard; peptide; 12 AA.

XX AAB82817;  
 XX 12-NOV-2001 (first entry)  
 XX Rabbit low density lipoprotein binding protein 2 (LBP-2) peptide.

XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX Oryctolagus cuniculus.

XX WO200164874-A2.  
 XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.  
 XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 14(c); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues  
 CC 354-365 of novel rabbit low density lipoprotein binding protein 2 (LBP-2,  
 CC see AAB82807). Rabbit LBP-2 is an example of claimed LBP polypeptides of  
 CC the invention that are capable of binding to native and methylated low  
 CC density lipoproteins (LDLs). Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs.

CC Polypeptides having the present amino acid sequence are among those  
 CC claimed. Methods of determining if an animal is at risk for  
 CC atherosclerosis, and methods for evaluating an agent for use in treating  
 CC atherosclerosis, are also claimed, as are pharmaceutical  
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
 CC compositions. Note: the present sequence is not shown in the  
 CC specification but is derived from the rabbit LBP-2 sequence given in  
 CC figure 2A (see AAB82807)

XX Sequence 12 AA;

Query Match 100.0%; Score 54; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.0043;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11  
 |||||  
 DB 2 VSEGSSEVPESD 12

RESULT 3

AAB82810  
 ID AAB82810 standard; peptide; 26 AA.

XX AAB82810;

XX 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 2 (LBP-2) peptide.

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 14(a); Page; 143pp; English.

XX The present sequence is that of a peptide comprising amino acid residues  
 CC 329-354 of novel human low density lipoprotein binding protein 2 (LBP-2,  
 CC see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of  
 CC the invention that are capable of binding to native and methylated low  
 CC density lipoproteins (LDLs). Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs.

CC Polypeptides having the present amino acid sequence are among those  
 CC claimed. Methods of determining if an animal is at risk for  
 CC atherosclerosis, and methods for evaluating an agent for use in treating  
 CC atherosclerosis, are also claimed, as are pharmaceutical  
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
 CC compositions. Note: the present sequence is not shown in the  
 CC specification but is derived from the human LBP-2 sequence given in  
 CC figure 7A (see AAB82806)

XX Sequence 26 AA;

Query Match 100.0%; Score 54; DB 4; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 0.01;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11  
 |||||  
 DB 16 VSEGSSEVPESD 26

RESULT 4

AAB82816  
 ID AAB82816 standard; peptide; 28 AA.

XX AAB82816;

```

XX 12-NOV-2001 (first entry)
DT
XX Rabbit low density lipoprotein binding protein 2 (LBP-2) peptide.
DE
XX
XX
XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX
XX Oryctolagus cuniculus.
OS
XX
XX WO200164874-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 28-FEB-2001; 2001WO-US06356.
PF
XX
XX 02-MAR-2000; 2000US-00517849.
PR
XX 14-JUL-2000; 2000US-00616289.
PR
XX (BOST-) BOSTON HEART FOUND INC.
PA
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
PI WPI; 2001-565505/63.
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
PT
XX Claim 14(c); Page; 143pp; English.
PS
XX
XX The present sequence is that of a peptide comprising amino acid residues
CC 338-365 of novel rabbit low density lipoprotein binding protein 2 (LBP-2,
CC see AAB82807). Rabbit LBP-2 is an example of claimed LBP polypeptides of
CC the invention that are capable of binding to native and methylated low
CC density lipoproteins (LDLs). Also claimed are biologically active
CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well
CC as expression vectors, cells and methods of producing the LBPs.
CC Polypeptides having the present amino acid sequence are among those
CC claimed. Methods of determining if an animal is at risk for
CC atherosclerosis, methods for evaluating an agent for use in treating
CC atherosclerosis, and methods for treating a cell having an abnormality in
CC structure or metabolism of LBP are also claimed, as are pharmaceutical
CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
CC compositions. Note: the present sequence is not shown in the
CC specification but is derived from the rabbit LBP-2 sequence given in
CC figure 2A (see AAB82807)
XX
XX Sequence 28 AA;
SQ
Query Match 100.0%; Score 54; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
Db 18 VSEGSSEVPESD 28
|||||
|||||

RESULT 5
AAW49041
ID AAW49041 standard; protein; 217 AA.
XX
XX AAW49041;
AC
XX
XX 09-NOV-1998 (first entry)
DT
XX Human low density lipoprotein binding protein LBP-2.
DE
XX
XX Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.
XX
XX Homo sapiens.
OS
XX

FH Key Location/Qualifiers
FT Peptide 8..33 /note= "Claim 2"
FT Peptide 8..22 /note= "Claim 2"
FT Peptide 23..33 /note= "Claim 2"
FT Peptide 208..217 /note= "Claim 2"
XX
XX WO9823282-A1.
PN
XX
XX 04-JUN-1998.
PD
XX
XX 26-NOV-1997; 97WO-US021857.
PF
XX
XX 27-NOV-1996; 96US-0031930P.
PR
XX 03-JUN-1997; 97US-0048547P.
PR
XX (BOST-) BOSTON HEART FOUND INC.
PA
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
PI WPI; 1998-322455/28.
XX N-PSDB; AAV32838.
DR
XX
XX Nucleic acid encoding low density lipoprotein binding proteins and
PT related vectors - transformed cells, proteins, and modulators of binding,
PT useful for treatment and diagnosis of atherosclerosis and for identifying
PT subjects at risk.
XX
XX Claim 1; Fig 7; 47pp; English.
PS
XX
XX This polypeptide comprises novel human low density lipoprotein (LDL)
CC binding protein LBP-2 that is capable of binding both native and methyl
CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see
CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human
CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP
CC metabolism or structure is diagnostic of a risk for atherosclerosis. The
CC invention provides: methods for determining if an animal is at risk for
CC atherosclerosis (e.g. for prenatal screening); methods for treating
CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to
CC bind LDL and thereby prevent formation of atherosclerotic plaque; and
CC methods for treating a cell having an abnormality in LBP structure or
CC metabolism. Pharmaceutical and vaccine compositions are also provided, as
CC well as recombinant vectors and host cells used to produce recombinant
CC LBP
XX
XX Sequence 217 AA;
SQ
Query Match 100.0%; Score 54; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11
Db 23 VSEGSSEVPESD 33
|||||
|||||

RESULT 6
AAB82803
ID AAB82803 standard; protein; 217 AA.
XX
XX AAB82803;
AC
XX
XX 12-NOV-2001 (first entry)
DT
XX Human low density lipoprotein binding protein 2 (LBP-2).
DE
XX
XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;
KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.
XX
XX Homo sapiens.
OS

```



PT polypeptides encoded by these genes, useful in the prevention, treatment  
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
 PT neurological diseases.

PS Claim 11; Page 1112; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic;  
 CC antiproliferative; antiviral; antiallergic; hepatotropic; antidiabetic;  
 CC antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;  
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and  
 CC protein sequences are used in the diagnosis of cancer, particularly  
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
 CC and agonists may also be used in the diagnosis, prevention and treatment  
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC cardiovascular disorders such as myocardial ischaemias; wound healing;  
 CC neurological diseases such as cerebral anoxia and epilepsy; and  
 CC infectious diseases

XX SQ Sequence 241 AA;

Query Match 100.0%; Score 54; DB 3; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 0.11; Indels 0; Gaps 0;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSSEVPESD 11  
 |||||

Db 47 VSEGSSEVPESD 57

RESULT 9  
 AAB82800

ID AAB82800 standard; protein; 252 AA.

AC AAB82800;

DT 12-NOV-2001 (first entry)

DE Rabbit low density lipoprotein binding protein 3 (LBP-3).

XX Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Oryctolagus cuniculus.

OS WO200164874-A2.

PN 07-SEP-2001.

PD 28-FEB-2001; 2001WO-US006356.

PF 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SM, Arjona AA;

XX WPI; 2001-565505/63.

DR N-PSDB; AAB26490.

XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(d); Fig 4; 143pp; English.

XX

CC The present sequence is that of a partial sequence of novel rabbit low  
 CC density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is  
 CC deduced from an isolated cDNA clone (see AAB26490). A full-length  
 CC sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an  
 CC example of claimed polypeptides of the invention, termed LBP3, that are  
 CC capable of binding to native and methylated low density lipoproteins.  
 CC Also claimed are biologically active fragments and analogues of LBP3,  
 CC polynucleotides encoding LBP3, as well as expression vectors, cells and  
 CC methods of producing the LBP3. Methods of determining if an animal is at  
 CC risk for atherosclerosis, methods for evaluating an agent for use in  
 CC treating atherosclerosis, and methods for treating a cell having an  
 CC abnormality in structure or metabolism of LBP are also claimed, as are  
 CC pharmaceutical compositions comprising an LBP polypeptide or nucleic  
 CC acid, and vaccine compositions

XX SQ Sequence 252 AA;

Query Match 100.0%; Score 54; DB 4; Length 252;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEGSSEVPESD 11  
 |||||

Db 57 VSEGSSEVPESD 67

RESULT 10

AAB49038

ID AAB49038 standard; protein; 317 AA.

XX AC AAB49038;

DT 09-NOV-1998 (first entry)

DE Rabbit low density lipoprotein binding protein LBP-2.

XX Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
 KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.

XX Oryctolagus cuniculus.

XX Key Location/Qualifiers

FT Misc-difference 10 /note= "encoded by TAG"

FT Protein 66..317 /note= "Claim 1"

FT Protein 86..317 /note= "Claim 1"

FT Peptide 105..132 /note= "Claim 2"

FT Peptide 105..120 /note= "Claim 2"

FT Peptide 121..132 /note= "Claim 2"

FT Peptide 211..220 /note= "Claim 2"

XX WO9823282-A1.

XX 04-JUN-1998.

XX 26-NOV-1997; 97WO-US021857.

XX 27-NOV-1996; 96US-0031930P.

XX 03-JUN-1997; 97US-0048547P.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SM, Arjona AA;

XX WPI; 1998-322455/28.

DR N-PSDB; AAB32835.

XX

PT Nucleic acid encoding low density lipoprotein binding proteins and  
 PT related vectors - transformed cells, proteins, and modulators of binding,  
 PT useful for treatment and diagnosis of atherosclerosis and for identifying  
 PT subjects at risk.

XX Claim 1; Fig 2; 47pp; English.

XX This polypeptide comprises novel rabbit low density lipoprotein (LDL)  
 CC binding protein LBP-2 that is capable of binding both native and methyl  
 CC LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA  
 CC (see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and  
 CC human LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of  
 CC LBP metabolism or structure is diagnostic of a risk for atherosclerosis.  
 CC The invention provides methods for determining if an animal is at risk  
 CC for atherosclerosis (e.g. for prenatal screening); methods for treating  
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to  
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and  
 CC methods for treating a cell having an abnormality in LBP structure or  
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as  
 CC well as recombinant vectors and host cells used to produce recombinant  
 CC LBP

XX Sequence 317 AA;

Query Match 100.0%; Score 54; DB 2; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 0; Indels 0;

QY 1 VSEGSSEVPESD 11  
 DB 122 VSEGSSEVPESD 132

RESULT 11

AAB82798

ID AAB82798 standard; protein; 317 AA.

XX AAB82798;

DT 12-NOV-2001 (first entry)

DE Rabbit low density lipoprotein binding protein 2 (LBP-2).

XX Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Oryctolagus cuniculus.

XX Key Location/Qualifiers

FT Misc-difference 10 /note= "encoded by TAG"

FT WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAH26488.

XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(b); Fig 2b; 143pp; English.

CC The present sequence is that of a partial sequence of novel rabbit low  
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
 CC deduced from an isolated cDNA clone (see AAH26488). Full-length rabbit  
 CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed  
 CC polypeptides of the invention, termed LBPs, that are capable of binding  
 CC to native and methylated low density lipoproteins. Also claimed are  
 CC biologically active fragments and analogues of LBPs, polynucleotides  
 CC encoding LBPs, as well as expression vectors, cells and methods of  
 CC producing the LBPs. Methods of determining if an animal is at risk for  
 CC atherosclerosis, methods for evaluating an agent for use in treating  
 CC atherosclerosis, and methods for treating a cell having an abnormality in  
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
 CC compositions

XX Sequence 317 AA;

Query Match 100.0%; Score 54; DB 4; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 0.15; Mismatches 0; Gaps 0;  
 Matches 11; Conservative 0; Indels 0;

QY 1 VSEGSSEVPESD 11  
 DB 122 VSEGSSEVPESD 132

RESULT 12

AAB82806

ID AAB82806 standard; protein; 538 AA.

XX AAB82806;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2).

XX Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAH26499.

XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein  
 CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the  
 CC coding region of isolated genomic DNA (see AAH26499). It differs from the  
 CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the  
 CC presence of an additional 321 amino acids at the N-terminus (the cDNA  
 CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP  
 CC polypeptides of the invention that are capable of binding to native and  
 CC methylated low density lipoproteins. Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs.  
 CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529

CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of  
 CC determining if an animal is at risk for atherosclerosis, methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed  
 SQ Sequence 538 AA;

Query Match 100.0%; Score 54; DB 4; Length 538;

Best Local Similarity 100.0%; Pred. No. 0.27; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11

|||||

Db 344 VSEGSSEVPESD 354

RESULT 13

AAB82807

ID AAB82807 standard; protein; 550 AA.

XX

AC AAB82807;

XX

DT 12-NOV-2001 (first entry)

XX

DE Rabbit low density lipoprotein binding protein 2 (LBP-2).

XX

KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;

XX

KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX

OS Oryctolagus cuniculus.

XX

PN WO200164874-A2.

XX

PD 07-SEP-2001.

XX

PF 28-FEB-2001; 2001WO-US006356.

XX

PR 02-MAR-2000; 2000US-00517849.

XX

PR 14-JUL-2000; 2000US-00616289.

XX

PA (BOST-) BOSTON HEART FOUND INC.

XX

PI Lees AM, Lees RS, Law SW, Arjona AA;

XX

DR WPI; 2001-555505/63.

XX

DR N-PSDB; AAH26500.

XX

PT New isolated low density lipoprotein binding polypeptide for treating,

XX

PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX

PS Claim 13(k); Fig 2a; 143pp; English.

XX

CC The present sequence is that of a full-length sequence of novel rabbit

CC low density lipoprotein binding protein 2 (LBP-2). The amino acid

CC sequence is deduced from an isolated cDNA clone (see AAH26500). Rabbit

CC LBP-2 is an example of claimed polypeptides of the invention, termed

CC LBP-2, that are capable of binding to native and methylated low density

CC lipoproteins. Also claimed are biologically active fragments and

CC analogues of LBPs, polynucleotides encoding LBPs, as well as expression

CC vectors, cells and methods of producing the LBPs. Polypeptides having

CC amino acid residues 338-353, 338-365, 354-365 or 444-453 (see AAB82815-

CC 18) of the present sequence are claimed. Methods of determining if an

CC animal is at risk for atherosclerosis, methods for evaluating an agent

CC for use in treating atherosclerosis, and methods for treating a cell

CC having an abnormality in structure or metabolism of LBP are also claimed,

CC as are pharmaceutical compositions comprising an LBP polypeptide or

CC nucleic acid, and vaccine compositions

SQ Sequence 550 AA;

Query Match 100.0%; Score 54; DB 4; Length 550;

Best Local Similarity 100.0%; Pred. No. 0.28; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11

|||||

Db 355 VSEGSSEVPESD 365

RESULT 14

ABM85225

ID ABM85225 standard; protein; 759 AA.

XX

AC ABM85225;

XX

DT 18-NOV-2004 (first entry)

XX

DE Human protein sequence hCP34885.

XX

XX Cytostatic; carcinoma; lymphoma; cancer; human.

XX

OS Homo sapiens.

XX

PN WO2003073826-A2.

XX

PD 12-SEP-2003.

XX

PF 28-FEB-2003; 2003WO-US006235.

XX

PR 01-MAR-2002; 2002US-00087192.

XX

PA (SAGR-) SAGRES DISCOVERY.

XX

PI Morris DW;

XX

DR WPI; 2003-328604/31.

XX

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX

PS Claim 5; SEQ ID NO 90; Opp; English.

XX

CC The present invention relates to novel DNA and protein sequences which

CC are associated with carcinomas. The sequences are useful for: (i) for

CC screening drug candidates; (ii) for screening of bioactive agent capable

CC of binding to Carcinoma Associated protein (CAP); (iii) for screening of

CC a bioactive agent capable of modulating the activity of CAP; (iv) for

CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating

CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for

CC determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of

CC carcinoma including lymphoma. The present sequence is one such CAP. Note:

CC This patent is an equivalent to basic patent US2002182586A1, for which no

CC sequence data was published

XX

SQ Sequence 759 AA;

Query Match 70.4%; Score 38; DB 7; Length 759;

Best Local Similarity 77.8%; Pred. No. 3.7e+02;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPE 9

|:|:|:|

Db 666 VTEGSQVPE 674

RESULT 15

ADN24118

ID ADN24118 standard; protein; 808 AA.

XX

AC ADN24118;

XX

Search completed: September 20, 2005, 12:34:45  
Job time : 7.54345 secs

DT 02-DEC-2004 (first entry)  
DE Bacterial polypeptide #6771.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
XX US2003233675-A1.  
PN  
XX  
PD 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
PF  
XX 21-FEB-2002; 2002US-0360039P.  
PR  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
PI WPI; 2004-061375/06.  
XX  
DR  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 6771; 122pp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 808 AA;

Query Match 70.4%; Score 38; DB 8; Length 808;  
Best Local Similarity 77.8%; Pred. No. 4e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 EGSEVPESD 11  
Db 338 EGSETPKSD 346



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 1.66952 Seconds  
(without alignments)  
491.841 Million cell updates/sec

Title: US-10-671-242-21  
Perfect score: 54  
Sequence: 1 VSEGSBPESD 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	11	3	US-08-979-608A-21
2	54	100.0	11	4	US-09-517-849-21
3	54	100.0	11	4	US-09-616-289-21
4	54	100.0	12	3	US-08-979-608A-27
5	54	100.0	12	4	US-09-517-849-27
6	54	100.0	12	4	US-09-616-289-27
7	54	100.0	26	3	US-08-979-608A-20
8	54	100.0	26	4	US-09-517-849-20
9	54	100.0	26	4	US-09-616-289-20
10	54	100.0	28	3	US-08-979-608A-26
11	54	100.0	28	4	US-09-517-849-26
12	54	100.0	28	4	US-09-616-289-26
13	54	100.0	217	3	US-08-979-608A-7
14	54	100.0	217	4	US-09-517-849-7
15	54	100.0	217	4	US-09-616-289-7
16	54	100.0	232	3	US-08-979-608A-3
17	54	100.0	232	4	US-09-517-849-3
18	54	100.0	232	4	US-09-616-289-3
19	54	100.0	252	3	US-08-979-608A-4
20	54	100.0	252	4	US-09-517-849-4
21	54	100.0	252	4	US-09-616-289-4
22	54	100.0	317	3	US-08-979-608A-2
23	54	100.0	317	4	US-09-517-849-2
24	54	100.0	317	4	US-09-616-289-2
25	54	100.0	538	4	US-09-616-289-43
26	54	100.0	550	4	US-09-616-289-47
27	36	66.7	277	4	US-09-949-016-6068

28	36	66.7	363	4	US-09-107-532A-5091	Sequence 5091, Ap
29	36	66.7	746	4	US-09-248-796A-20280	Sequence 20280, A
30	35.5	65.7	755	4	US-09-134-000C-6691	Sequence 6691, Ap
31	35	64.8	263	4	US-09-134-000C-5721	Sequence 5721, Ap
32	35	64.8	287	4	US-09-489-039A-8878	Sequence 8878, Ap
33	35	64.8	415	4	US-09-252-991A-23602	Sequence 23602, A
34	35	64.8	456	4	US-09-902-540-11227	Sequence 11227, A
35	34	63.0	97	4	US-09-621-976-6690	Sequence 6690, Ap
36	34	63.0	121	3	US-09-134-001C-3834	Sequence 3834, Ap
37	34	63.0	173	4	US-08-956-171B-5218	Sequence 5218, Ap
38	34	63.0	173	4	US-08-781-986A-5218	Sequence 5218, Ap
39	34	63.0	373	4	US-09-710-279-372	Sequence 372, Appl
40	34	63.0	591	4	US-09-602-787A-24	Sequence 24, Appl
41	34	63.0	712	4	US-09-708-426-9	Sequence 9, Appli
42	34	63.0	857	2	US-08-659-251-2	Sequence 2, Appli
43	34	63.0	857	3	US-09-256-490-2	Sequence 2, Appli
44	34	63.0	857	5	PCT-US96-11445-2	Sequence 2, Appli
45	34	63.0	1247	1	US-08-472-934-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-08-979-608A-21  
; Sequence 21, Application US/08979608A  
; Patent No. 6355451

GENERAL INFORMATION:

APPLICANT: Lees, Ann M. S.  
Lees, Robert S.  
Law, Simon W.  
Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,608A

FILING DATE: 26-Nov. 6355451-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,547

FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930

FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 21:

US-08-979-608A-21

Query Match 100.0%; Score 54; DB 3; Length 11;

```
Best Local Similarity 100.0%; Pred. No. 0.00097; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 2
US-09-517-849-21
; Sequence 21, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Law, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-517-849-21

Query Match 100.0%; Score 54; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 3
US-09-616-289-21
; Sequence 21, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

Best Local Similarity 100.0%; Score 54; DB 4; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
ATHEROSCLEROSIS
FILE REFERENCE: 10797-004001
CURRENT APPLICATION NUMBER: US/09/616,289
CURRENT FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/517,849
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 08/979,608
PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: US 60/031,930
PRIOR FILING DATE: 1996-11-27
PRIOR APPLICATION NUMBER: US 60/048,547
PRIOR FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 53
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-09-616-289-21

Query Match 100.0%; Score 54; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11
Db 1 VSEGSVPESD 11

RESULT 4
US-08-979-608A-27
; Sequence 27, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
```

;  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-08-979-608A-27

Query Match 100.0%; Score 54; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGEVPSD 11  
Db |||||

## RESULT 5

US-09-517-849-27

; Sequence 27, Application US/09517849

; Patent No. 6605588

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish &amp; Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/517,849

; FILING DATE: 02-Mar-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-003001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-517-849-27

Query Match 100.0%; Score 54; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGEVPSD 11  
Db |||||

## RESULT 6

US-09-616-289-27  
; Sequence 27, Application US/09616289  
; Patent No. 6632923

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; APPLICANT: Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING

; ATHEROSCLEROSIS

; FILE REFERENCE: 10797-004001

; CURRENT APPLICATION NUMBER: US/09/616,289

; CURRENT FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 09/517,849

; PRIOR FILING DATE: 2000-03-02

; PRIOR APPLICATION NUMBER: US 08/979,608

; PRIOR FILING DATE: 1997-11-26

; PRIOR APPLICATION NUMBER: US 60/031,930

; PRIOR FILING DATE: 1996-11-27

; PRIOR APPLICATION NUMBER: US 60/048,547

; PRIOR FILING DATE: 1997-06-03

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 27

; LENGTH: 12

; TYPE: PRT

; ORGANISM: Oryctolagus cuniculus

; US-09-616-289-27

Query Match 100.0%; Score 54; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGEVPSD 11  
Db |||||

## RESULT 7

US-08-979-608A-20

; Sequence 20, Application US/08979608A

; Patent No. 6355451

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish &amp; Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/979,608A

; FILING DATE: 26-No. 6355451-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,547

; FILING DATE: 03-JUN-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-979-608A-20

Query Match      100.0%; Score 54; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VSEGSVPESD 11
Db      16 VSEGSVPESD 26

RESULT 8
US-09-517-849-20
; Sequence 20, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20

Query Match      100.0%; Score 54; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
```

```
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VSEGSVPESD 11
Db      16 VSEGSVPESD 26

RESULT 9
US-09-616-289-20
; Sequence 20, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 26
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-616-289-20

Query Match      100.0%; Score 54; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VSEGSVPESD 11
Db      16 VSEGSVPESD 26

RESULT 10
US-08-979-608A-26
; Sequence 26, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-Nov-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-517-849-20

Query Match      100.0%; Score 54; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0026;
```

/ APPLICATION NUMBER: US/08/979,608A  
/ FILING DATE: 26-NO. 6355451-1997  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US 60/048,547  
/ FILING DATE: 03-JUN-1997  
/ APPLICATION NUMBER: US 60/031,930  
/ FILING DATE: 27-NOV-1996  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Myers, Louis  
/ REGISTRATION NUMBER: 35,965  
/ REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 617/542-5070  
/ TELEFAX: 617/542-8906  
/ INFORMATION FOR SEQ ID NO: 26:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 28 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-08-979-608A-26

Query Match 100.0%; Score 54; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11  
DB 18 VSEGSSEVPESD 28

RESULT 11  
US-09-517-849-26  
/ Sequence 26, Application US/09517849  
/ Patent No. 6605588  
/ GENERAL INFORMATION:  
/ APPLICANT: Lees, Ann M.  
/ Lees, Robert S.  
/ Law, Simon W.  
/ Arjona, Anibal A.  
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
/ TREATING ATHEROSCLEROSIS  
/ NUMBER OF SEQUENCES: 42  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Fish & Richardson P.C.  
/ STREET: 225 Franklin Street  
/ CITY: Boston  
/ STATE: MA  
/ COUNTRY: USA  
/ ZIP: 02110-2804  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Diskette  
/ COMPUTER: IBM Compatible  
/ OPERATING SYSTEM: DOS  
/ SOFTWARE: FastSEQ for Windows Version 2.0  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/517,849  
/ FILING DATE: 02-Mar-2000  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: 08/979,608  
/ FILING DATE: 26-NOV-1997  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Myers, Louis  
/ REGISTRATION NUMBER: 35,965  
/ REFERENCE/DOCKET NUMBER: 10797-003001  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 617/542-5070  
/ TELEFAX: 617/542-8906  
/ INFORMATION FOR SEQ ID NO: 26:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 28 amino acids

/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
US-09-517-849-26

Query Match 100.0%; Score 54; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11  
DB 18 VSEGSSEVPESD 28

RESULT 12  
US-09-616-289-26  
/ Sequence 26, Application US/09616289  
/ Patent No. 6632923  
/ GENERAL INFORMATION:  
/ APPLICANT: Lees, Ann M.  
/ APPLICANT: Lees, Robert S.  
/ APPLICANT: Law, Simon W.  
/ APPLICANT: Arjona, Anibal A.  
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
/ PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
/ ATHEROSCLEROSIS  
/ FILE REFERENCE: 10797-004001  
/ CURRENT APPLICATION NUMBER: US/09/616,289  
/ CURRENT FILING DATE: 2000-07-14  
/ PRIOR APPLICATION NUMBER: US 09/517,849  
/ PRIOR FILING DATE: 2000-03-02  
/ PRIOR APPLICATION NUMBER: US 08/979,608  
/ PRIOR FILING DATE: 1997-11-26  
/ PRIOR APPLICATION NUMBER: US 60/031,930  
/ PRIOR FILING DATE: 1996-11-27  
/ PRIOR APPLICATION NUMBER: US 60/048,547  
/ PRIOR FILING DATE: 1997-06-03  
/ NUMBER OF SEQ ID NOS: 53  
/ SOFTWARE: FastSEQ for Windows Version 4.0  
/ SEQ ID NO 26  
/ LENGTH: 28  
/ TYPE: PRT  
/ ORGANISM: Oryctolagus cuniculus  
US-09-616-289-26

Query Match 100.0%; Score 54; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSSEVPESD 11  
DB 18 VSEGSSEVPESD 28

RESULT 13  
US-08-979-608A-7  
/ Sequence 7, Application US/08979608A  
/ Patent No. 6355451  
/ GENERAL INFORMATION:  
/ APPLICANT: Lees, Ann M.  
/ Lees, Robert S.  
/ Law, Simon W.  
/ Arjona, Anibal A.  
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
/ TREATING ATHEROSCLEROSIS  
/ NUMBER OF SEQUENCES: 42  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Fish & Richardson P.C.  
/ STREET: 225 Franklin Street  
/ CITY: Boston  
/ STATE: MA

```
/
/ COUNTRY: USA
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/979,608A
/ FILING DATE: 26-NOV-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/048,547
/ FILING DATE: 03-JUN-1997
/ APPLICATION NUMBER: US 60/031,930
/ FILING DATE: 27-NOV-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 217 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-979-608A-7

Query Match 100.0%; Score 54; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSSEVPESD 11
DB 23 VSEGSSEVPESD 33

RESULT 14
US-09-517-849-7
/ Sequence 7, Application US/09517849
/ Patent No. 6605588
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ Lees, Robert S.
/ Law, Simon W.
/ Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
/ BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
/ TREATING ATHEROSCLEROSIS
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/517,849
/ FILING DATE: 02-Mar-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/979,608
/ FILING DATE: 26-NOV-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Myers, Louis
```

```
/
/ REGISTRATION NUMBER: 35,965
/ REFERENCE/DOCKET NUMBER: 10797-003001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 217 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-517-849-7

Query Match 100.0%; Score 54; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSSEVPESD 11
DB 23 VSEGSSEVPESD 33

RESULT 15
US-09-616-289-7
/ Sequence 7, Application US/09616289
/ Patent No. 6632923
/ GENERAL INFORMATION:
/ APPLICANT: Lees, Ann M.
/ Lees, Robert S.
/ Law, Simon W.
/ APPLICANT: Arjona, Anibal A.
/ TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
/ PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
/ ATHEROSCLEROSIS
/ FILE REFERENCE: 10797-004001
/ CURRENT APPLICATION NUMBER: US/09/616,289
/ CURRENT FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 09/517,849
/ PRIOR FILING DATE: 2000-03-02
/ PRIOR APPLICATION NUMBER: US 08/979,608
/ PRIOR FILING DATE: 1997-11-26
/ PRIOR APPLICATION NUMBER: US 60/031,930
/ PRIOR FILING DATE: 1996-11-27
/ PRIOR APPLICATION NUMBER: US 60/048,547
/ PRIOR FILING DATE: 1997-06-03
/ NUMBER OF SEQ ID NOS: 53
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 217
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-616-289-7

Query Match 100.0%; Score 54; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.029; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 1 VSEGSSEVPESD 11
DB 23 VSEGSSEVPESD 33

Search completed: September 20, 2005, 12:45:30
Job time : 2.66952 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 1.23868 Seconds  
(without alignments)  
854.447 Million cell updates/sec

Title: US-10-671-242-21

Perfect score: 54

Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	75.9	244	2 F69066	conserved hypother
2	41	75.9	587	2 E84808	hypothetical prote
3	38	70.4	139	2 AF0221	conserved hypother
4	38	70.4	808	2 T16564	hypothetical prote
5	38	70.4	906	2 A43817	transforming prote
6	38	70.4	1661	2 A83695	hypothetical prote
7	37	68.5	258	2 D95284	probable [imported
8	37	68.5	437	2 S57776	cysteine proteinas
9	37	68.5	447	2 T46146	hypothetical prote
10	37	68.5	831	2 A70363	mannose-1-phosphat
11	37	68.5	2218	2 B84683	hypothetical prote
12	36	66.7	75	2 E87604	conserved hypother
13	36	66.7	204	2 T33362	hypothetical prote
14	36	66.7	496	2 T52112	deoxyribodipyrimid
15	36	66.7	723	2 S51788	malate synthase (E
16	36	66.7	1877	2 T21861	hypothetical prote
17	35	64.8	216	2 H69443	conserved hypother
18	35	64.8	316	2 E89973	hypothetical prote
19	35	64.8	1173	2 T51892	hypothetical prote
20	35	64.8	1883	2 G82875	hypothetical prote
21	35	64.8	2301	2 T02323	nodulin-like prote
22	35	64.8	6420	2 T30283	polyketide synthas
23	34.5	63.9	341	2 S43586	F26F3.4 protein (c
24	34	63.0	164	2 E89955	hypothetical prote
25	34	63.0	183	2 C97832	alpha-(1,3)-fucosyl
26	34	63.0	281	2 JC4295	heat-shock protein
27	34	63.0	332	2 A56123	streptogrisin D (E
28	34	63.0	513	2 A96265	hypothetical prote
29	34	63.0	513	2 AH3019	sigma 54 dependent

30	34	63.0	551	1 VGNZP2	cell fusion glycop
31	34	63.0	551	1 VGNZPG	cell fusion glycop
32	34	63.0	552	2 T44954	flagella-related p
33	34	63.0	590	2 T42202	probable acyl-CoA
34	34	63.0	590	2 T42206	probable acyl-CoA
35	34	63.0	613	2 T24662	hypothetical prote
36	34	63.0	638	2 T24661	hypothetical prote
37	34	63.0	639	2 B95945	probable glycosyl
38	34	63.0	728	2 AF3299	malate synthase (E
39	34	63.0	819	2 C84615	hypothetical prote
40	34	63.0	877	2 T03098	p97 protein - Toxo
41	34	63.0	896	2 B43817	transforming prote
42	34	63.0	1022	2 T24663	hypothetical prote
43	34	63.0	1039	2 T43678	tetratricopeptide
44	34	63.0	1039	2 T38447	tetratricopeptide
45	34	63.0	1129	2 T42732	A-kinase anchoring

ALIGNMENTS

RESULT 1

F69066

Conserved hypothetical protein MTH1498 - Methanobacterium thermoautotrophicum (strain Del

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: F69066

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F

i; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.,

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: F69066

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-244 <MTH>

A;Cross-references: UNIPROT:O27542; GB:AE000910; GB:AE000666; NID:g2622610; PIDN:AAB8597;

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1498

A;Start codon: GTG

C;Superfamily: conserved hypothetical protein MTH593

Query Match 75.9%; Score 41; DB 2; Length 244;  
Best Local Similarity 54.5%; Pred. No. 3.6;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSEGSVPESD 11

Db 231 ITEGEIPEED 241

RESULT 2

E84808

hypothetical protein At2g38720 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C;Accession: E84808

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: E84808

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-587 <STO>

A;Cross-references: UNIPROT:Q9ZVJ3; GB:AE002093; NID:g3786000; PIDN:AAC67346.1; GSPDB:GN

C;Genetics:

A;Gene: At2g38720

A;Map position: 2





A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A;Reference number: A96039; MUID:21368234; PMID:11474104

C;Contents: annotation

C;Genetics:

A;Gene: SMA0335

A;Genome: plasmid

C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 68.5%; Score 37; DB 2; Length 258;

Best Local Similarity 60.0%; Pred. No. 23;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SEGSEVPESD 11

Db 103 TDGAEPVEDD 112

RESULT 8

S57776

cysteine proteinase (EC 3.4.22.-) - clove pink (fragment)

C;Species: Dianthus caryophyllus (clove pink)

C;Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004

C;Accession: S57776

R;Jones, M.L.; Larsen, P.B.; Woodson, W.R.

Plant Mol. Biol. 28, 505-512, 1995

A;Title: Ethylene-regulated expression of a carnation cysteine proteinase during flower

A;Reference number: S57776; MUID:95359407; PMID:7632919

A;Accession: S57776

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-427 <JON>

A;Cross-references: UNIPROT:Q43423; EMBL:U17135; NID:G595985; PIDN:AAA79915.1; PID:G5959

C;Superfamily: papain

C;Keywords: cysteine proteinase; hydrolase

P;119,255,276/Active site: Cys, His, Asn #status predicted

Query Match 68.5%; Score 37; DB 2; Length 427;

Best Local Similarity 70.0%; Pred. No. 39;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VSEGSEVPES 10

Db 89 VKEGDELPS 98

RESULT 9

T46146

hypothetical protein T3A5.70 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C;Accession: T46146

R;Bloeker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, P.; Sala

submitted to the Protein Sequence Database, December 1999

A;Reference number: 223024

A;Accession: T46146

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-447 <BLO>

A;Cross-references: UNIPROT:Q9SCQ7; EMBL:AL132979

A;Experimental source: cultivar Columbia; BAC clone T3A5

C;Genetics:

A;Map position: 3

A;Introns: 369/3

A;Note: T3A5.70

Query Match 68.5%; Score 37; DB 2; Length 447;

Best Local Similarity 70.0%; Pred. No. 41;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SEGSEVPESD 11

Db 156 AEGNERPESD 165

RESULT 10

A70363

mannose-1-phosphate guanyltransferase - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C;Accession: A70363

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

V.

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196686; PMID:9537320

A;Accession: A70363

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-831 <AQF>

A;Cross-references: UNIPROT:O66933; GB:AE000704; NID:G2983301; PIDN:AAC06893.1; PID:G2983

A;Experimental source: strain VFS

C;Genetics:

A;Gene: mpg

Query Match 68.5%; Score 37; DB 2; Length 831;

Best Local Similarity 60.0%; Pred. No. 81;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VSEGSEVPES 10

Db 254 VEEGTPEIPEN 263

RESULT 11

B84683

hypothetical protein At2g28300 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C;Accession: B84683

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: B84683

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-2218 <STO>

A;Cross-references: GB:AE002093; NID:G4803953; PIDN:AAD29825.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g28300

A;Map position: 2

Query Match 68.5%; Score 37; DB 2; Length 2218;

Best Local Similarity 70.0%; Pred. No. 2.4e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VSEGSEVPES 10

Db 2071 VTEGSELPSS 2080

RESULT 12

E87604

conserved hypothetical protein CC2869 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C;Accession: E87604

R;Nierman, W.C.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: E87604

```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <STO>
A;Cross-references: UNIPROT:Q9A4G5; GB:AE005673; NID:gl3424485; PIDN:AAK24833.1; GSPDB:G
C;Genetics:
A;Gene: CC2869

Query Match          66.7%; Score 36; DB 2; Length 75;
Best Local Similarity 54.5%; Pred. No. 9.2;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSEGSVEPESD 11
Db 58 LSEGRKIPEPD 68

RESULT 13
T33362
hypothetical protein F16G10.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33362
R;Gattung, S.; Scheet, P.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid F16G10.
A;Reference number: 221329
A;Accession: T33362
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-204 <GAT>
A;Cross-references: UNIPROT:O76595; EMBL:AF077537; PIDN:AAC26278.1; GSPDB:GNO00020; CESP:
C;Genetics:
A;Experimental source: strain Bristol N2; clone F16G10
A;Gene: CESP:F16G10.10
A;Map position: 2
A;Introns: 56/1; 134/3

Query Match          66.7%; Score 36; DB 2; Length 204;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGSEVPESD 11
Db 120 EGSELPKAD 128

RESULT 14
T52112
deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) [validated] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52112
R;Ahmad, M.; Jarillo, J.A.; Cashmore, A.R.
submitted to the EMBL Data Library, March 1998
A;Reference number: 225961
A;Accession: T52112
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-496 <AHM>
A;Cross-references: UNIPROT:O24374; EMBL:AF053365; PIDN:AAC08008.1
C;Genetics:
A;Gene: PHR1
C;Function:
A;Description: EC 4.1.99.3 [validated, MUID:97215604]
C;Keywords: carbon-carbon lyase

Query Match          66.7%; Score 36; DB 2; Length 496;
Best Local Similarity 77.8%; Pred. No. 72;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VSEGSVEPVE 9
Db 218 VREGAEVPE 226

```

## RESULT 15

```

S51788
malate synthase (EC 4.1.3.2) isoenzyme G - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 23-Aug-1995 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S51788; F65083
R;Molina, I.; Pellicer, M.T.; Badia, J.; Aguilar, J.; Baldona, L.
Eur. J. Biochem. 224, 541-548, 1994
A;Title: Molecular characterization of Escherichia coli malate synthase G. Differentiation
A;Reference number: S51788; MUID:95010032; PMID:7925370
A;Accession: S51788
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-723 <MOL>
A;Cross-references: UNIPROT:P37330; EMBL:X74547; NID:g517246; PIDN:CRA52639.1; PID:g51724
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65083
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-723 <BLAT>
A;Cross-references: GB:AE000380; GB:U00096; NID:gl789344; PIDN:AAC76012.1; PID:gl789348;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: glcB
C;Keywords: carbon-carbon lyase; coenzyme A; oxo-acid-lyase

Query Match          66.7%; Score 36; DB 2; Length 723;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SEGSVEPVE 9
Db 352 SEGNHPE 359

Search completed: September 20, 2005, 12:43:22
Job time : 3.23868 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 5.58751 Seconds  
(without alignments)  
1008.117 Million cell updates/sec

Title: US-10-671-242-21  
Perfect score: 54  
Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	236	2 Q96IM4	O6p03 homo sapien
2	54	100.0	285	2 Q6P0R3	O6p03 homo sapien
3	54	100.0	295	2 Q6P1S7	O6p17 homo sapien
4	54	100.0	538	2 Q6SPF0	O6sp0 homo sapien
5	54	100.0	550	2 Q6SPF0	O6sp9 oryctolagus
6	42	77.8	714	2 Q6CC87	O6cc87 yarrowia li
7	41	75.9	244	2 Q27542	O27542 methanobact
8	41	75.9	587	2 Q9ZVJ3	O9zv13 arabidopsis
9	40	74.1	150	2 Q6L129	O6l129 picophilus
10	40	74.1	2972	2 Q8FR52	O8ff52 corynebacte
11	39	72.2	465	2 Q6SXA3	O6sxa3 oryza sativ
12	38	70.4	139	2 Q8ZFA0	O8zfa0 yersinia pe
13	38	70.4	192	2 Q67LF8	O67lf8 symbiobacte
14	38	70.4	501	1 PTGI RAT	O62969 rattus norv
15	38	70.4	677	2 Q9GYN2	Q9gyn2 caenorhabdi
16	38	70.4	798	2 Q6CUR8	Q6cur8 kluyveromyc
17	38	70.4	906	1 CBL HUMAN	P22681 homo sapien
18	38	70.4	1528	2 Q75B25	Q75b25 ashbya goss
19	38	70.4	1661	2 Q9KFW0	Q9kfw0 bacillus ha
20	38	70.4	1678	2 Q924C5	Q924c5 mus musculu
21	37	68.5	258	2 Q930L5	Q930l5 rhizobium m
22	37	68.5	269	2 Q8L9D7	Q8l9d7 arabidopsis
23	37	68.5	404	2 Q8KE30	Q8ke30 chlorobium
24	37	68.5	427	2 Q434Z3	Q434z3 dianthus ca
25	37	68.5	447	2 Q9SCQ7	Q9scq7 arabidopsis
26	37	68.5	499	2 Q71B18	Q71b18 stellaria l
27	37	68.5	574	2 Q8NS53	Q8ns53 corynebacte
28	37	68.5	581	2 Q6YSM8	Q6ysm8 mus musculu
29	37	68.5	675	2 Q7S816	Q7s816 neurospora
30	37	68.5	768	2 Q64PM0	Q64pm0 bacteroides
31	37	68.5	831	2 Q66933	Q66933 aquifex aeo

32	37	68.5	832	2 Q9N4C1	Q9n4c1 caenorhabdi
33	37	68.5	957	2 Q94GY5	Q94gy5 oryza sativ
34	37	68.5	1512	2 Q8CDL2	Q8cdl2 mus musculu
35	37	68.5	1911	2 Q8OU10	Q8ou10 mus musculu
36	37	68.5	3313	2 Q9W3W6	Q9w3w6 drosophila
37	37	68.5	3571	2 Q9SL27	Q9sl27 arabidopsis
38	37	68.5	3574	2 Q9AUB4	Q9aub4 arabidopsis
39	37	68.5	4685	2 Q93HJ2	Q93hj2 streptomyce
40	36	66.7	75	2 Q9A4G5	Q9a4g5 caulobacter
41	36	66.7	164	2 Q8WMM9	Q8wmm9 ovis aries
42	36	66.7	169	2 Q8FTU3	Q8ftu3 homo sapien
43	36	66.7	204	2 Q76595	Q76595 caenorhabdi
44	36	66.7	248	2 Q7RZW8	Q7rzw8 neurospora
45	36	66.7	278	1 IOD3_HUMAN	P55073 homo sapien

#### ALIGNMENTS

RESULT 1  
Q96IM4 PRELIMINARY; PRT; 236 AA.  
AC Q96IM4  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE LOC90378 protein (Hypothetical protein) (Fragment).  
GN Name=LOC90378;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Director MGC Project;  
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC007384; AAH07384.2; -;  
DR EMBL; BC080588; AAH80588.1; -;  
DR HSSP; P39769; 1KW4.  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR010993; SAM\_homology.  
DR Pfam; PF00536; SAM 1; 1.  
DR SMART; SM00454; SAM; 1.

```

DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 54; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 42 VSEGSEVPESD 52
|||||

RESULT 2
Q6POR3 PRELIMINARY; PRT; 285 AA.
AC Q6POR3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Udgin T.B., Toshikiyuki S., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AAH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAFF1F901 CRC64;

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC90378 protein (fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Udgin T.B., Toshikiyuki S., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 101 VSEGSEVPESD 111
|||||

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atherin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RC SEQUENCE FROM N.A.
RA TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY453840; AAR24087.1; -.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM_1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS0105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 54; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 344 VSEGSEVPESD 354
|||||
|

RESULT 5
Q6SPE9 PRELIMINARY; PRT; 550 AA.
AC Q6SPE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acherin.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RC SEQUENCE FROM N.A.
RA TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY453841; AAR24088.1; -.
DR InterPro: IPR001660; SAM.
DR Pfam: PF00536; SAM_1; 1.
DR SMART: SM00454; SAM; 1.
DR PROSITE: PS0105; SAM_DOMAIN; 1.
SQ SEQUENCE 550 AA; 57019 MW; AF1CF2B780D879A7 CRC64;

Query Match 100.0%; Score 54; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
Db 355 VSEGSEVPESD 365
|||||
|

RESULT 6
Q6CC87 PRELIMINARY; PRT; 714 AA.
AC Q6CC87;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome C of strain CLIB99 of Yarrowia
DE lipolytica.
GN ORFNames=YALI0C11495g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RC SEQUENCE FROM N.A.
RA Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

```

```

RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: CR382129; CAG82035.1; -.
DR GO: GO:0004835; F:ubulin-tyrosine ligase activity; IEA.
DR GO: GO:0006464; P:protein modification; IEA.
DR InterPro: IPR002828; SurE.
DR InterPro: IPR004344; Tub_tyr_lygase.
DR Pfam: PF01975; SurE; 1.
DR Pfam: PF03133; TTL; 1.
DR ProDom: PD005378; SurE; 1.
DR TIGRFAMs: TIGR00087; surE; 1.
SQ SEQUENCE 714 AA; 80159 MW; 26DCA372AD90FE63 CRC64;

Query Match 77.8%; Score 42; DB 2; Length 714;
Best Local Similarity 88.9%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGSEVPESD 11
Db 700 EGSEVPESD 708
|||||
|

RESULT 7
O27542 PRELIMINARY; PRT; 244 AA.
ID O27542;
AC O27542;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved protein.
GN OrderedLocusNames=MT1498;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL: AE000910; AAB85973.1; -.
DR PIR: F69066; F69066.
DR InterPro: IPR010624; KaIC.
DR Pfam: PF06745; KaIC; 1.
KW Complete proteome.
SQ SEQUENCE 244 AA; 27537 MW; A93F15849B5CCC3B CRC64;

```

```
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 150 AA; 17477 MW; 0F81FBE1FC98EA49 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 150;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVEPESD 11
:|||||:
Db 231 ITGESEIPEED 241

RESULT 8
Q9ZVJ3 PRELIMINARY; PRT; 587 AA.
AC Q9ZVJ3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein At2g38720.
GN Name=At2g38720;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Renning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005499; AAC67346.1; -.
DR PIR; E84808; E84808.
DR InterPro; IPR007145; MAP65_ASE1.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF03999; MAP65_ASE1; 1.
KW Hypothetical protein.
SQ SEQUENCE 587 AA; 67137 MW; 5A3A708A72119425 CRC64;

Query Match 75.9%; Score 41; DB 2; Length 587;
Best Local Similarity 81.8%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSEGSVEPESD 11
:|||||:
Db 137 VSSGSEVDES 147

RESULT 9
Q6L129 PRELIMINARY; PRT; 150 AA.
AC Q6L129;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PT00738;
OS Picrophilus torridus.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Picrophilaceae; Picrophilus.
OX NCBI_TaxID=82076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 9790 / ATCC 700027;
RX PubMed=15184674; DOI=10.1073/pnas.0401356101;
RA Fuetteler O., Angelov A., Liesegang H., Gottschalk G., Schleper C.,
RA Scheper B., Dock C., Antranikian G., Liebl W.;
RT "Genome sequence of Picrophilus torridus and its implications for life
RT around pH 0."
RL Proc. Natl. Acad. Sci. U.S.A. 101:9091-9096(2004).
DR EMBL; AE017261; AAT43323.1; -.

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 150 AA; 17477 MW; 0F81FBE1FC98EA49 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 150;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSVEPES 9
:|||||:
Db 36 VSEGSKVPE 44

RESULT 10
Q8FR52 PRELIMINARY; PRT; 2972 AA.
AC Q8FR52;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fatty-acid synthase I (EC 2.3.1.85).
GN Name=fasA; OrderedLocusNames=CE0913;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RA "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens."
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP05217; BAC17723.1; -.
DR HSSP; O32472; L106.
DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR003965; Fatty_acid_synth.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; Ketoacyl-synt_C; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
KW Acyltransferase; Complete proteome; Transferase.
SQ SEQUENCE 2972 AA; 316363 MW; C0EFFF5DDFC123C9 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 2972;
Best Local Similarity 70.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SEGSVEPESD 11
:|||||:
Db 2400 AEGTEVPEED 2409

RESULT 11
Q65XA3 PRELIMINARY; PRT; 465 AA.
ID Q65XA3
AC Q65XA3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
```

DE Putative MAP kinase phosphatase.  
GN Names=OJ1654\_B10.15;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,  
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,  
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
RA Wu H.-P., Shaw J.-F.;  
RT "Oryza sativa BAC OJ1654\_B10 genomic sequence.";  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC108504; AAU44084.1; -;  
KW Kinase.  
SQ SEQUENCE 465 AA; 50818 MW; E8D433E00006AEF9 CRC64;  
Query Match 72.2%; Score 39; DB 2; Length 465;  
Best Local Similarity 88.9%; Pred. No. 1.4e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VSEGSEVPE 9  
| | | | |  
DB 75 VREGSEVPE 83  
RESULT 12  
Q82FA0 PRELIMINARY; PRT; 139 AA.  
ID Q82FA0 Q74DW8; O7CHY6;  
AC Q82FA0 Q74DW8; O7CHY6;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
DE Hypothetical protein YP01817.  
GN OrderedLocusNames=Yp1576, YP01817, Y2490;  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586350; DOI=10.1038/35097083;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIMS / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RX DOI=10.1128/JB.184.16.4601-4611.2002;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Petherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM.";  
RL J. Bacteriol. 184:4601-4611(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=91001 / Biovar Mediaevalis;  
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,  
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,

RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
RA Yang R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ414150; CAC90634.1; -;  
DR EMBL; AB013852; AAM86047.1; -;  
DR EMBL; AB017132; AAS61810.1; -;  
DR PIR; AF0221; AF0221.  
DR PFAM; PF03681; UPF0150; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 139 AA; 15549 MW; A608044F830649B1 CRC64;  
Query Match 70.4%; Score 38; DB 2; Length 139;  
Best Local Similarity 70.0%; Pred. No. 53;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VSEGSEVPE 10  
| | | | |  
DB 52 VSEGIEIPE 61  
RESULT 13  
Q67LF8 PRELIMINARY; PRT; 192 AA.  
ID Q67LF8 Q57LF8;  
AC Q57LF8 Q57LF8;  
DT 25-OCT-2004 (T-EMBLrel. 28, Created)  
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
DE Metalloendopeptidase-related protein.  
GN ORFNames=STH2503;  
OS Symbiobacterium thermophilum.  
OC Bacteria; Actinobacteria; Symbiobacterium.  
OX NCBI\_TaxID=2734;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IAM14863;  
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsui T.,  
RA Morimura K., Ikeda H., Hattori M., Beppu T.;  
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium thermophilum.";  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP005840; BAD41488.1; -;  
DR InterPro; IPR011053; Hybrid motif.  
DR InterPro; IPR002886; Peptidase\_M23B.  
DR PFAM; PF01551; Peptidase\_M23; 1.  
SQ SEQUENCE 192 AA; 20535 MW; C0ABEC13A4F282D0 CRC64;  
Query Match 70.4%; Score 38; DB 2; Length 192;  
Best Local Similarity 63.6%; Pred. No. 76;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 VSEGSEVPE 11  
| | | | |  
DB 142 VSEGDRVPOGD 152  
RESULT 14  
PTGI RAT STANDARD; PRT; 501 AA.  
ID PTGI RAT STANDARD; PRT; 501 AA.  
AC Q62369;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Prostacyclin synthase (EC 5.3.99.4) (Prostaglandin 12 synthase).  
GN Name=Ptgis; Synonyms=Cyp8;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Geraci M.W., Gao B., Shepherd D., Moore M., Vernon J., Miller Y.E.,  
RA Voelkel N.F.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Catalyzes the isomerization of prostaglandin H2 to  
CC prostacyclin (= prostaglandin I2) (By similarity).  
CC -!- CATALYTIC ACTIVITY: (5Z,13E)-(15S)-9-alpha,11-alpha-epidioxy-15-  
CC hydroxyprosta-5,13-dienoate = (5Z,13E)-(15S)-6,9-alpha-epoxy-11-  
CC alpha,15-dihydroxyprosta-5,13-dienoate.  
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane. May be  
CC anchored to the membrane via a single transmembrane domain.  
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U53855; AAB02322.1; -.  
CC RGD; 3438; Ptgis.  
CC InterPro; IPR001128; Cytochrome\_P450.  
CC InterPro; IPR002403; EP450IV.  
CC Pfam; PF00067; P450; 2.  
CC PRINTS; PR00465; EP450IV.  
CC PRINTS; PR00385; P450.  
CC PROSITE; PS00086; CYTOCHROME\_P450; FALSE NEG.  
CC Endoplasmic reticulum; Heme; Isomerase; Prostaglandin biosynthesis;  
CC Transmembrane.  
CC TRANSMEM 1 21 Potential.  
CC FT METAL 442 442 Iron (heme axial ligand) (By similarity).  
CC SQ SEQUENCE 501 AA; 57127 MW; D2D85ADP3C464863 CRC64;  
  
Query Match 70.4%; Score 38; DB 1; Length 501;  
Best Local Similarity 72.7%; Pred. No. 2.3e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 VSEGEVPEPSD 11  
||| :||| |  
DB 468 VSEDTVEPFD 478  
  
RESULT 15  
Q9GYN2  
ID Q9GYN2 PRELIMINARY; PRT; 677 AA.  
AC Q9GYN2;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Patterned expression site protein 22.  
GN Names=pep-22; ORFNames=K05B2.5;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium. ";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Chissee S.;  
RT "The sequence of C. elegans cosmid K05B2. ";  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]

RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilson R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilson R.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U29379; AAF99981.3; -.  
DR WormBase; WBGene00003986; K05B2.5.  
DR WormPep; K05B2.5; CE34052.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR PROSITE; PS50850; MFS; 1.  
SQ SEQUENCE 677 AA; 73704 MW; 478B6ADE2DF389BE CRC64;  
  
Query Match 70.4%; Score 38; DB 2; Length 677;  
Best Local Similarity 77.8%; Pred. No. 3.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 EGSEVPESD 11  
||| :||| |  
DB 269 EGSETPKSD 277  
  
Search completed: September 20, 2005, 12:41:46  
Job time : 7.58751 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 6.07222 Seconds  
(without alignments)  
733.538 Million cell updates/sec

Title: US-10-671-242-21

Perfect score: 54

Sequence: 1 VSEGSVPESD 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/PCT PUBCOMB.pcp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB.pcp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06 PUB.pcp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06 PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07 NEW PUB.pcp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08 NEW PUB.pcp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08 PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09 PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09 PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09 PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09 NEW PUB.pcp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10A PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C PUBCOMB.pcp.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10C PUBCOMB.pcp.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10C NEW PUB.pcp.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A PUBCOMB.pcp.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11 NEW PUB.pcp.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60 NEW PUB.pcp.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60 PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	11	9	US-09-962-055-21
2	54	100.0	11	9	US-09-976-740-21
3	54	100.0	11	13	US-10-023-529-21
4	54	100.0	11	13	US-10-023-529-21
5	54	100.0	11	15	US-10-023-529-21
6	54	100.0	11	15	US-10-023-529-21
7	54	100.0	12	9	US-09-962-055-27
8	54	100.0	12	9	US-09-976-740-27
9	54	100.0	12	13	US-10-023-529-27
10	54	100.0	12	13	US-10-023-529-27
11	54	100.0	12	15	US-10-023-529-27

12	54	100.0	12	15	US-10-671-242-27	Sequence 27, Appl
13	54	100.0	26	9	US-09-962-055-20	Sequence 20, Appl
14	54	100.0	26	9	US-09-976-740-20	Sequence 20, Appl
15	54	100.0	26	13	US-10-023-529-20	Sequence 20, Appl
16	54	100.0	26	13	US-10-023-529-20	Sequence 20, Appl
17	54	100.0	26	15	US-10-023-529-20	Sequence 20, Appl
18	54	100.0	26	15	US-10-023-529-20	Sequence 20, Appl
19	54	100.0	28	9	US-09-962-055-26	Sequence 26, Appl
20	54	100.0	28	9	US-09-976-740-26	Sequence 26, Appl
21	54	100.0	28	13	US-10-023-529-26	Sequence 26, Appl
22	54	100.0	28	13	US-10-023-529-26	Sequence 26, Appl
23	54	100.0	28	15	US-10-023-529-26	Sequence 26, Appl
24	54	100.0	28	15	US-10-023-529-26	Sequence 26, Appl
25	54	100.0	217	9	US-09-962-055-7	Sequence 7, Appl
26	54	100.0	217	9	US-09-976-740-7	Sequence 7, Appl
27	54	100.0	217	13	US-10-023-529-7	Sequence 7, Appl
28	54	100.0	217	13	US-10-023-529-7	Sequence 7, Appl
29	54	100.0	217	15	US-10-023-529-7	Sequence 7, Appl
30	54	100.0	217	15	US-10-023-529-7	Sequence 7, Appl
31	54	100.0	232	9	US-09-962-055-3	Sequence 3, Appl
32	54	100.0	232	9	US-09-976-740-3	Sequence 3, Appl
33	54	100.0	232	13	US-10-023-529-3	Sequence 3, Appl
34	54	100.0	232	13	US-10-023-529-3	Sequence 3, Appl
35	54	100.0	232	15	US-10-023-529-3	Sequence 3, Appl
36	54	100.0	232	15	US-10-023-529-3	Sequence 3, Appl
37	54	100.0	241	9	US-09-925-298-665	Sequence 665, App
38	54	100.0	241	14	US-10-102-806-665	Sequence 665, App
39	54	100.0	252	9	US-09-962-055-4	Sequence 4, Appl
40	54	100.0	252	9	US-09-976-740-4	Sequence 4, Appl
41	54	100.0	252	13	US-10-023-529-4	Sequence 4, Appl
42	54	100.0	252	13	US-10-023-529-4	Sequence 4, Appl
43	54	100.0	252	15	US-10-023-529-4	Sequence 4, Appl
44	54	100.0	252	15	US-10-023-529-4	Sequence 4, Appl
45	54	100.0	317	9	US-09-962-055-2	Sequence 2, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-962-055-21

; Sequence 21, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

APPLICANT: Lees, Ann M.

Lees, Robert S.

Artjona, Anibal A.

Law, Simon W.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/962,055

FILING DATE: 24-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/979,608

FILING DATE: 26-NOV-1997

APPLICATION NUMBER: US 60/031,930

FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:

```
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-962-055-21

Query Match      100.0%; Score 54; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSSEVPESD 11
Db      1 VSEGSSEVPESD 11

RESULT 2
US-09-976-740-21
; Sequence 21, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-21

Query Match      100.0%; Score 54; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSSEVPESD 11
Db      1 VSEGSSEVPESD 11

RESULT 3
US-10-023-529-21
; Sequence 21, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-21

Query Match      100.0%; Score 54; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSSEVPESD 11
Db      1 VSEGSSEVPESD 11

RESULT 4
US-10-023-523-21
; Sequence 21, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-21

Query Match      100.0%; Score 54; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGSSEVPESD 11
Db      1 VSEGSSEVPESD 11
```

RESULT 5  
US-10-616-187-21  
; Sequence 21, Application US/10616187  
; Publication No. US20040013668A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/10/616,187  
; PRIOR FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: US/09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-616-187-21

Query Match 100.0%; Score 54; DB 15; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11  
|||||

Db 1 VSEGSEVPESD 11

RESULT 6  
US-10-671-242-21  
; Sequence 21, Application US/10671242  
; Publication No. US20040040049A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/10/671,242  
; PRIOR FILING DATE: 2003-09-24  
; PRIOR APPLICATION NUMBER: US/09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 11

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-671-242-21

Query Match 100.0%; Score 54; DB 15; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0058;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11  
|||||

Db 1 VSEGSEVPESD 11

RESULT 7  
US-09-962-055-27  
; Sequence 27, Application US/09962055  
; Patent No. US20020052033A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/962,055  
; FILING DATE: 24-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/979,608  
; FILING DATE: 26-NOV-1997  
; APPLICATION NUMBER: US 60/031,930  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-962-055-27

Query Match 100.0%; Score 54; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0064;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11  
|||||

Db 2 VSEGSEVPESD 12

RESULT 8  
US-09-976-740-27

```

; Sequence 27, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-976-740-27

```

```

Query Match      100.0%; Score 54; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 VSEGEVPESD 11
Db      2 VSEGEVPESD 12

```

## RESULT 9

```

US-10-023-529-27
; Sequence 27, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-529-27

```

```

Query Match      100.0%; Score 54; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;

```

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGEVPESD 11
Db      2 VSEGEVPESD 12

```

## RESULT 10

```

US-10-023-523-27
; Sequence 27, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR FILING DATE: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-023-523-27

```

```

Query Match      100.0%; Score 54; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 VSEGEVPESD 11
Db      2 VSEGEVPESD 12

```

## RESULT 11

```

US-10-616-187-27
; Sequence 27, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR FILING DATE: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547

```

```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-616-187-27

Query Match      100.0%; Score 54; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
   |||||
Db 2 VSEGSEVPESD 12

RESULT 12
US-10-671-242-27
; Sequence 27, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671.242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-671-242-27

Query Match      100.0%; Score 54; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
   |||||
Db 2 VSEGSEVPESD 12

RESULT 13
US-09-962-055-20
; Sequence 20, Application US/09962055
; Patent No. US20020052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,055
; FILING DATE: 24-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-962-055-20

Query Match      100.0%; Score 54; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSEGSEVPESD 11
   |||||
Db 16 VSEGSEVPESD 26

RESULT 14
US-09-976-740-20
; Sequence 20, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-976-740-20

Query Match      100.0%; Score 54; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGEVPSD 11
Db      16 VSEGEVPSD 26

RESULT 15
US-10-023-529-20
; Sequence 20, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-20

Query Match      100.0%; Score 54; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSEGEVPSD 11
Db      16 VSEGEVPSD 26

Search completed: September 20, 2005, 12:53:06
Job time : 6.07222 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 11:58:28 ; Search time 5.94859 Seconds  
(without alignments)  
650.171 Million cell updates/sec

Title: US-10-671-242-22

Perfect score: 58

Sequence: 1 EDDPDGFLG 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	10	4	AAB82812 Human low
2	58	100.0	132	7	AD59158 Human Pro
3	58	100.0	217	2	AAM49041 Human low
4	58	100.0	217	4	AAB82803 Human low
5	58	100.0	241	3	AAB88957 Breat an
6	58	100.0	538	4	AAB82806 Human low
7	54	93.1	232	4	AAB82799 Rabbit lo
8	54	93.1	252	4	AAB82800 Rabbit lo
9	54	93.1	317	2	AAM49038 Rabbit lo
10	54	93.1	317	4	AAB82798 Rabbit lo
11	54	93.1	550	4	AAB82807 Rabbit lo
12	44	75.9	945	4	AAM51653 Arabidops
13	43	74.1	100	3	AAG02194 Human sec
14	43	74.1	346	5	ABB06721 Human wit
15	43	74.1	804	4	ABB69136 Drosophil
16	41	70.7	140	5	ABP31168 Human ORF
17	41	70.7	354	4	AAM41146 Propionib
18	41	70.7	354	6	ABM40665 Propionib
19	41	70.7	354	7	AD31306 Human nov
20	41	70.7	354	8	ADR46143 Human pro
21	41	70.7	354	8	ADR46142 Human pro
22	41	70.7	435	8	ADQ65987 Novel hum
23	41	70.7	510	8	ADS29627 Bacterial
24	39	67.2	14	8	ADI00477 Influenza
25	39	67.2	14	8	ADI56752 Flu virus

## ALIGNMENTS

## RESULT 1

AAB82812

ID AAB82812 standard; peptide; 10 AA.

XX AAB82812;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2) peptide.

KW Low density lipoprotein binding protein 2; LBP-2; LDL; human; atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

XX Homo sapiens.

XX WO200164874-A2.

PD 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US006356.

PR 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

(BOST-) BOSTON HEART FOUND INC.

PI Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

PT New isolated low density lipoprotein binding polypeptide for treating, diagnosing and/or identifying therapeutic agents for atherosclerosis.

Claim 14(a); Page; 143pp; English.

The present sequence is that of a peptide comprising amino acid residues 529-538 of novel human low density lipoprotein binding protein 2 (LBP-2, see AAB82806). Human LBP-2 is an example of claimed LBP polypeptides of the invention that are capable of binding to native and methylated low density lipoproteins (LDLs). Also claimed are biologically active fragments and analogues of LBPs, polynucleotides encoding LBPs, as well as expression vectors, cells and methods of producing the LBPs. Polypeptides having the present amino acid sequence are among those claimed. Methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP are also claimed, as are pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine

Abp02064 Human ORF  
Abu44831 Protein e  
Adh39670 Streptomy  
Adh39812 Streptomy  
Adn20818 Bacterial  
Adal9453 Mouse LAT  
Aam25741 Human pro  
Aae06606 Human pro  
Ade07899 Novel pro  
Aam94615 Human rep  
Abm84853 Human dia  
Aay01459 Polypepti  
Abo23396 Amino aci  
Adr20783 Human sec  
Aay41706 Human PRO  
Aay01458 Polypepti  
Aab33463 Human PRO  
Aay94975 Human sec  
Aab44262 Human PRO  
Aab24021 Human PRO

CC compositions. Note: the present sequence is not shown in the  
CC specification but is derived from the human LBP-2 sequence given in  
CC figure 7A (see AAB2806)

XX  
SQ Sequence 10 AA;

Query Match 100.0%; Score 58; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0053;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
|||||  
Db 1 EDDDPDGFGLG 10

RESULT 2  
ADE59158  
ID ADE59158 standard; protein; 132 AA.  
XX  
AC ADE59158;  
XX  
DT 29-JAN-2004 (first entry)  
DE Human Protein XP\_031299, SEQ ID NO 5049.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
XX WO2003016475-A2.  
PN  
XX  
XX 27-FEB-2003.  
XX  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-033347P.  
XX  
XX (GEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; XP\_031299.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 132 AA;

Query Match 100.0%; Score 58; DB 7; Length 132;  
Best Local Similarity 100.0%; Pred. No. 0.074;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
|||||  
Db 123 EDDDPDGFGLG 132

RESULT 3  
AAW49041  
ID AAW49041 standard; protein; 217 AA.  
XX  
AC AAW49041;  
XX  
DT 09-NOV-1998 (first entry)  
DE Human low density lipoprotein binding protein LBP-2.  
XX  
KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH FT Peptide 8..33 /note= "Claim 2"  
FT FT Peptide 8..22 /note= "Claim 2"  
FT FT Peptide 23..33 /note= "Claim 2"  
FT FT Peptide 208..217 /note= "Claim 2"  
XX  
XX WO98223282-A1.  
PN  
XX  
XX 04-JUN-1998.  
PD  
XX  
XX 26-NOV-1997; 97WO-US021857.  
PF  
XX  
XX 27-NOV-1996; 96US-0031930P.  
PR 03-JUN-1997; 97US-0048547P.  
XX  
XX (BOST-) BOSTON HEART FOUND INC.  
XX  
XX Lees AM, Lees RS, Law SW, Arjona AA;  
PI  
XX  
XX WPI; 1998-322455/28.  
DR N-P8DB; AAV32838.  
XX  
XX Nucleic acid encoding low density lipoprotein binding proteins and  
PT related vectors - transformed cells, proteins, and modulators of binding,  
PT useful for treatment, and diagnosis of atherosclerosis and for identifying  
PT subjects at risk.  
XX  
XX Claim 1; Fig 7; 47pp; English.

CC This polypeptide comprises novel human low density lipoprotein (LDL)  
CC binding protein LBP-2 that is capable of binding both native and methyl  
CC LDL. Its amino acid sequence was deduced from an isolated cDNA clone (see  
CC AAV32838). cDNA clones (see AAV32834-39) and encoded rabbit and human  
CC LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP



CC Metabolism or structure is diagnostic of a risk for atherosclerosis. The  
 CC invention provides: methods for determining if an animal is at risk for  
 CC atherosclerosis (e.g. for prenatal screening); methods for treating  
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to  
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and  
 CC methods for treating a cell having an abnormality in LBP structure or  
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as  
 CC well as recombinant vectors and host cells used to produce recombinant  
 CC LBP

SQ Sequence 217 AA;

Query Match 100.0%; Score 58; DB 2; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
 |||||  
 Db 208 EDDDPDGFGLG 217

RESULT 4

AAB82803  
 ID AAB82803 standard; protein; 217 AA.

AC AAB82803;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 2 (LBP-2).

KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.

OS Homo sapiens.

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

XX N-PSDB; AAB26494.

PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

PS Claim 13 (g); Fig 7B; 143pp; English.

XX The present sequence is that of the N-terminal portion of novel human low  
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
 CC deduced from an isolated partial cDNA clone (see AAB26494). A full-length  
 CC sequence is given in AAB82806. Human LBP-2 is an example of claimed LBP  
 CC polypeptides of the invention that are capable of binding to native and  
 CC methylated low density lipoproteins. Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs. Methods  
 CC of determining if an animal is at risk for atherosclerosis, methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed

XX Sequence 217 AA;

Query Match 100.0%; Score 58; DB 4; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
 |||||  
 Db 208 EDDDPDGFGLG 217

RESULT 5

AAB58957  
 ID AAB58957 standard; protein; 241 AA.

XX AAB58957;

DT 27-MAR-2001 (first entry)

XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 665.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
 KW neutropenic; neutropenic; antiviral; antiallergic; hepatotropic;  
 KW antidiabetic; antiinflammatory; antitumor; antitumor; antitumor;  
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
 KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005881.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX N-PSDB; AAF21860.

PT New human breast and ovarian cancer associated gene sequences and the  
 PT polypeptides encoded by these genes, useful in the prevention, treatment  
 PT and diagnosis of cancer, immune disorders, cardiovascular disorders and  
 PT neurological diseases.

PS Claim 11; Page 1112; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterization of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic; immunosuppressive; neutropenic;  
 CC neutropenic; antiviral; antiallergic; hepatotropic; antidiabetic;  
 CC antiinflammatory; antitumor; antitumor; antitumor; antibacterial;  
 CC antifungal; antiparasitic and cardiant activity. The polynucleotide and  
 CC protein sequences are used in the diagnosis of cancer, particularly  
 CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists  
 CC and agonists may also be used in the diagnosis, prevention and treatment  
 CC of immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC cardiovascular disorders such as myocardial ischaemia; wound healing;  
 CC neurological diseases such as cerebral anoxia and epilepsy; and  
 CC infectious diseases

XX Sequence 241 AA;

Query Match 100.0%; Score 58; DB 3; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 0.14; 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
 |||||  
 Db 232 EDDDPDGFGLG 241

RESULT 6  
 AAB82806  
 ID AAB82806 standard; protein; 538 AA.  
 XX  
 AC AAB82806;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Human low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; human;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 PA (BOST-) BOSTON HEART FOUND INC.  
 XX  
 PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX  
 DR WPI; 2001-565505/63.  
 DR N-PSDB; AAH26499.  
 XX  
 PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 PS Claim 13(j); Fig 7A; 143pp; English.

XX The present sequence is that of novel human low density lipoprotein  
 CC binding protein 2 (LBP-2). The amino acid sequence was deduced from the  
 CC coding region of isolated genomic DNA (see AAH26499). It differs from the  
 CC sequence predicted from an LBP-2 cDNA clone (see AAB82803) by the  
 CC presence of an additional 321 amino acids at the N-terminus (the cDNA  
 CC clone is 5' truncated). Human LBP-2 is an example of claimed LBP  
 CC polypeptides of the invention that are capable of binding to native and  
 CC methylated low density lipoproteins. Also claimed are biologically active  
 CC fragments and analogues of LBPs, polynucleotides encoding LBPs, as well  
 CC as expression vectors, cells and methods of producing the LBPs.  
 CC Polypeptides having amino acid residues 329-343, 329-354, 344-354, or 529  
 CC -538 (see AAB82809-12) of the present sequence are claimed. Methods of  
 CC determining if an animal is at risk for atherosclerosis, methods for  
 CC evaluating an agent for use in treating atherosclerosis, and methods for  
 CC treating a cell having an abnormality in structure or metabolism of LBP  
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or  
 CC nucleic acid, and vaccine compositions, are also claimed  
 XX

Sequence 538 AA;  
 Query Match 100.0%; Score 58; DB 4; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
 |||||  
 Db 529 EDDDPDGFGLG 538

RESULT 7  
 AAB82799  
 ID AAB82799 standard; protein; 232 AA.  
 XX  
 AC AAB82799;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Rabbit low density lipoprotein binding protein 2 (LBP-2).  
 XX  
 KW Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 OS Oryctolagus cuniculus.  
 XX  
 PN WO200164874-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006356.  
 XX  
 PR 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 PA (BOST-) BOSTON HEART FOUND INC.  
 XX  
 PI Lees AM, Lees RS, Law SW, Arjona AA;  
 XX  
 DR WPI; 2001-565505/63.  
 DR N-PSDB; AAH26489.  
 XX  
 PT New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 PS Claim 13(c); Fig 3; 143pp; English.

XX The present sequence is that of a partial sequence of novel rabbit low  
 CC density lipoprotein binding protein 2 (LBP-2). The amino acid sequence is  
 CC deduced from an isolated cDNA clone (see AAH26489). Full-length rabbit  
 CC LBP-2 is given in AAB82807. Rabbit LBP-2 is an example of claimed  
 CC polypeptides of the invention, termed LBPs, that are capable of binding  
 CC to native and methylated low density lipoproteins. Also claimed are  
 CC biologically active fragments and analogues of LBPs, polynucleotides  
 CC encoding LBPs, as well as expression vectors, cells and methods of  
 CC producing the LBPs. Methods of determining if an animal is at risk for  
 CC atherosclerosis, methods for evaluating an agent for use in treating  
 CC atherosclerosis, and methods for treating a cell having an abnormality in  
 CC structure or metabolism of LBP are also claimed, as are pharmaceutical  
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine  
 CC compositions  
 XX

Sequence 232 AA;  
 Query Match 93.1%; Score 54; DB 4; Length 232;  
 Best Local Similarity 90.0%; Pred. No. 0.61;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
 |||||  
 Db 223 EDDDPDGFGLG 232

RESULT 8  
 AAB82800  
 ID AAB82800 standard; protein; 252 AA.  
 XX  
 AC AAB82800;  
 XX  
 DT 12-NOV-2001 (first entry)  
 XX  
 DE Rabbit low density lipoprotein binding protein 3 (LBP-3).

XX Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 XX Oryctolagus cuniculus.  
 XX  
 XX W0200164874-A2.  
 XX  
 XX 07-SEP-2001.  
 XX  
 XX 28-FEB-2001; 2001WO-US006356.  
 XX  
 XX 02-MAR-2000; 2000US-00517849.  
 PR 14-JUL-2000; 2000US-00616289.  
 XX  
 XX (BOST-) BOSTON HEART FOUND INC.  
 XX  
 XX Lees AM, Lees RS, Law SW, Arjona AA;  
 XX WPI; 2001-565505/63.  
 DR N-PSDB; AAH26490.  
 XX  
 XX New isolated low density lipoprotein binding polypeptide for treating,  
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.  
 XX  
 XX Claim 13 (d); Fig 4; 143pp; English.  
 XX  
 XX The present sequence is that of a partial sequence of novel rabbit low  
 CC density lipoprotein binding protein 3 (LBP-3). The amino acid sequence is  
 CC deduced from an isolated cDNA clone (see AAH26490). A full-length  
 CC sequence for rabbit LBP-3 is given in AAB82801. Rabbit LBP-3 is an  
 CC example of claimed polypeptides of the invention, termed LBPs, that are  
 CC capable of binding to native and methylated low density lipoproteins.  
 CC Also claimed are biologically active fragments and analogues of LBPs.  
 CC polynucleotides encoding LBPs, as well as expression vectors, cells and  
 CC methods of producing the LBPs. Methods for determining if an animal is at  
 CC risk for atherosclerosis, methods for evaluating an agent for use in  
 CC treating atherosclerosis, and methods for treating a cell having an  
 CC abnormality in structure or metabolism of LBP are also claimed, as are  
 CC pharmaceutical compositions comprising an LBP polypeptide or nucleic  
 CC acid, and vaccine compositions  
 XX  
 XX Sequence 252 AA;  
 SQ  
 Query Match 93.1%; Score 54; DB 4; Length 252;  
 Best Local Similarity 90.0%; Pred. No. 0.67;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EDDDPDGFLG 10  
 DB 243 EDDDPDGFLG 252  
 |||||:||||  
 RESULT 9  
 ID AAH49038  
 XX AAH49038 standard; protein; 317 AA.  
 AC  
 XX AAH49038;  
 XX  
 XX 09-NOV-1998 (first entry)  
 DT  
 XX  
 XX Rabbit low density lipoprotein binding protein LBP-2.  
 DE  
 DE Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;  
 KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine.  
 XX  
 XX Oryctolagus cuniculus.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH  
 FT Misc-difference 10  
 FT /note= "encoded by TAG"  
 FT Protein 66..317  
 FT /note= "Claim 1"

FT Protein 86..317  
 FT /note= "Claim 1"  
 FT Peptide 105..132  
 FT /note= "Claim 2"  
 FT Peptide 105..120  
 FT /note= "Claim 2"  
 FT Peptide 121..132  
 FT /note= "Claim 2"  
 FT Peptide 211..220  
 FT /note= "Claim 2"  
 XX  
 XX W09823282-A1.  
 PN  
 XX 04-JUN-1998.  
 PD  
 XX 26-NOV-1997; 97WO-US021857.  
 PF  
 XX 27-NOV-1996; 96US-0031930P.  
 PR 03-JUN-1997; 97US-0048547P.  
 XX  
 XX (BOST-) BOSTON HEART FOUND INC.  
 XX  
 XX Lees AM, Lees RS, Law SW, Arjona AA;  
 PI WPI; 1998-322455/28.  
 XX N-PSDB; AAV32835.  
 DR  
 XX Nucleic acid encoding low density lipoprotein binding proteins and  
 PT related vectors - transformed cells, proteins, and modulators of binding,  
 PT useful for treatment and diagnosis of atherosclerosis and for identifying  
 PT subjects at risk.  
 XX  
 XX Claim 1; Fig 2; 47pp; English.  
 PS  
 XX This polypeptide comprises novel rabbit low density lipoprotein (LDL)  
 CC binding protein LBP-2 that is capable of binding both native and methyl  
 CC LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA  
 CC (see AAV32835). cDNA clones (see AAV32834-39) and encoded rabbit and  
 CC human LBPs (see AAH49037-42) are claimed. An abnormality in an aspect of  
 CC LBP metabolism or structure is diagnostic of a risk for atherosclerosis.  
 CC The invention provides: methods for determining if an animal is at risk  
 CC for atherosclerosis (e.g. for prenatal screening); methods for treating  
 CC atherosclerosis (including gene therapy) using e.g. LBP polypeptides to  
 CC bind LDL and thereby prevent formation of atherosclerotic plaque; and  
 CC methods for treating a cell having an abnormality in LBP structure or  
 CC metabolism. Pharmaceutical and vaccine compositions are also provided, as  
 CC well as recombinant vectors and host cells used to produce recombinant  
 CC LBP  
 XX  
 XX Sequence 317 AA;  
 SQ  
 Query Match 93.1%; Score 54; DB 2; Length 317;  
 Best Local Similarity 90.0%; Pred. No. 0.84;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EDDDPDGFLG 10  
 DB 308 EDDDPDGFLG 317  
 |||||:||||  
 RESULT 10  
 ID AAB82798  
 XX AAB82798 standard; protein; 317 AA.  
 AC  
 XX AAB82798;  
 XX  
 XX 12-NOV-2001 (first entry)  
 DT  
 XX Rabbit low density lipoprotein binding protein 2 (LBP-2).  
 DE  
 DE Low density lipoprotein binding protein 2; LBP-2; LDL; rabbit;  
 KW atherosclerosis; antiarteriosclerotic; therapy; diagnosis; vaccine.  
 XX  
 XX





## RESULT 15

ABB69136  
 ID ABB69136 standard; protein; 804 AA.  
 XX  
 AC ABB69136;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 34200.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL13239.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 34200; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 804 AA;  
 Query Match 74.1%; Score 43; DB 4; Length 804;  
 Best Local Similarity 77.8%; Pred. NO. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 DDDPDGFLG 10  
 Db |||:||||  
 483 DDSPEGLG 491

Search completed: September 20, 2005, 12:34:47  
 Job time : 7.94859 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:14:43 ; Search time 1.51775 Seconds  
(without alignments)  
491.841 Million cell updates/sec

Title: US-10-671-242-22  
Perfect score: 58  
Sequence: 1 EDDDPDGFLG 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	10	3	US-08-979-608A-22
2	58	100.0	10	4	US-09-517-849-22
3	58	100.0	10	4	US-09-616-289-22
4	58	100.0	217	3	US-08-979-608A-7
5	58	100.0	217	4	US-09-517-849-7
6	58	100.0	217	4	US-09-616-289-7
7	58	100.0	538	4	US-09-616-289-43
8	54	93.1	232	3	US-08-979-608A-3
9	54	93.1	232	4	US-09-517-849-3
10	54	93.1	232	4	US-09-616-289-3
11	54	93.1	252	3	US-08-979-608A-4
12	54	93.1	252	4	US-09-517-849-4
13	54	93.1	252	4	US-09-616-289-4
14	54	93.1	317	3	US-08-979-608A-2
15	54	93.1	317	4	US-09-517-849-2
16	54	93.1	317	4	US-09-616-289-2
17	54	93.1	550	4	US-09-616-289-47
18	43	74.1	100	4	US-09-513-999C-6275
19	41	70.7	109	4	US-09-621-976-7055
20	38	65.5	328	4	US-09-902-540-13561
21	38	65.5	826	4	US-09-248-796A-14387
22	37	63.8	198	3	US-09-228-986-93
23	37	63.8	198	4	US-10-101-464A-93
24	37	63.8	211	4	US-10-101-464A-762
25	37	63.8	227	4	US-09-248-796A-21586
26	37	63.8	260	4	US-09-589-927-8
27	37	63.8	260	4	US-09-277-665-8

28	37	63.8	260	4	US-09-589-987-8	Sequence 8, Appli
29	37	63.8	455	2	US-08-738-172-4	Sequence 4, Appli
30	37	63.8	485	4	US-09-252-991A-17170	Sequence 17170, A
31	37	63.8	575	4	US-09-248-796A-14348	Sequence 14348, A
32	37	63.8	596	4	US-10-101-464A-889	Sequence 889, App
33	37	63.8	996	4	US-10-101-464A-933	Sequence 933, App
34	36	62.1	263	4	US-09-902-540-12633	Sequence 12633, A
35	36	62.1	354	4	US-09-949-016-10579	Sequence 10579, A
36	36	62.1	427	4	US-09-248-796A-19520	Sequence 19520, A
37	36	62.1	459	4	US-09-252-991A-24552	Sequence 24552, A
38	36	62.1	469	4	US-09-411-132A-5	Sequence 5, Appli
39	36	62.1	703	4	US-09-252-991A-20065	Sequence 20065, A
40	36	62.1	704	1	US-08-190-802A-62	Sequence 62, Appli
41	36	62.1	704	1	US-08-188-583-5	Sequence 5, Appli
42	36	62.1	704	1	US-08-646-715-5	Sequence 5, Appli
43	36	62.1	704	2	US-08-308-818-3	Sequence 3, Appli
44	36	62.1	704	3	US-08-477-346-62	Sequence 62, Appli
45	36	62.1	704	3	US-08-473-089-62	Sequence 62, Appli

ALIGNMENTS

RESULT 1  
US-08-979-608A-22  
; Sequence 22, Application US/08979608A  
; Patent No. 6355451  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M. S.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,608A  
FILING DATE: 26-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/048,547  
FILING DATE: 03-JUN-1997  
APPLICATION NUMBER: US 60/031,930  
FILING DATE: 27-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-08-979-608A-22

Query Match 100.0%; Score 58; DB 3; Length 10;

```
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 2
US-09-517-849-22
; Sequence 22, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/517,849
; FILING DATE: 02-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,608
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-517-849-22

Query Match 100.0%; Score 58; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 3
US-09-616-289-22
; Sequence 22, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
```

```
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-616-289-22

Query Match 100.0%; Score 58; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 4
US-08-979-608A-7
; Sequence 7, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-No. 6355451-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```



;  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-08-979-608A-7

Query Match 100.0%; Score 58; DB 3; Length 217;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
Db 208 EDDDPDGFGLG 217

## RESULT 5

US-09-517-849-7  
; Sequence 7, Application US/09517849  
; Patent No. 6605588  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/517,849

FILING DATE: 02-Mar-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/979,608

FILING DATE: 26-Nov-1997

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-003001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 217 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-517-849-7

Query Match 100.0%; Score 58; DB 4; Length 217;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
Db 208 EDDDPDGFGLG 217

## RESULT 6

0;



```
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Cryptolagus cuniculus
US-09-616-289-3

Query Match          93.1%; Score 54; DB 4; Length 232;
Best Local Similarity 90.0%; Pred. No. 0.08;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
    |||||.||||
Db 223 EDDDPDGFLG 232

RESULT 11
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match          93.1%; Score 54; DB 3; Length 252;
Best Local Similarity 90.0%; Pred. No. 0.088;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
    |||||.||||
Db 243 EDDDPDGFLG 252

RESULT 12
US-08-979-608A-4
; Sequence 4, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                   TREATING ATHEROSCLEROSIS
;
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,608A
; FILING DATE: 26-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/048,547
; FILING DATE: 03-JUN-1997
; APPLICATION NUMBER: US 60/031,930
; FILING DATE: 27-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-979-608A-4

Query Match          93.1%; Score 54; DB 3; Length 252;
Best Local Similarity 90.0%; Pred. No. 0.088;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
    |||||.||||
Db 243 EDDDPDGFLG 252

RESULT 13
US-09-616-289-4
; Sequence 4, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
;               Lees, Robert S.
;               Law, Simon W.
;               Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;                   PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;                   ATHEROSCLEROSIS
;
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
```

```

; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; US-09-616-289-4

```

Query Match	93.1%	Score 54;	DB 4;	Length 252;
Best Local Similarity	90.0%;	Pred.	No. 0.088;	
Matches 9; Conservative		1;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy 1 EDDDPGFLG 10  
|||:|  
Db 243 EDDDPGFLG 252

RESULT 14  
 US-08-979-608A-2  
 ; Sequence 2, Application US/08979608A  
 ; Patent No. 6355451  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lees, Ann M.  
 ; Lees, Robert S.  
 ; Law, Simon W.  
 ; Arjona, Anibal A.  
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
 ; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
 ; TREATING ATHEROSCLEROSIS  
 ;  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA

Query Match	93.1%	Score 54;	DB 3;	Length 317;
Best Local Similarity	90.0%;	Pred. No. 0.11;		

	Matches	9;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	EDDDPDCFLG	10							
		:								
Db	308	EDDDPDCFLG	317							

RESULT 15  
US-09-517-849-2  
; Sequence 2, Application US/09517849  
; Patent NO. 660588  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
;

Query Match	93.1%	Score 54;	DB 4;	Length 317;
Best Local Similarity	90.0%;	Pred. No. 0.11;		
Matches	9;	Conservative	1;	Mismatches
				0;
				Indels
				0;
				Gaps
				0;

Qy 1 EDDDPDGLG 10  
Dy 308 EDDDPGLG 317

Search completed: September 20, 2005, 12:45:31  
Job time : 2.51775 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:11:23 ; Search time 1.12607 Seconds  
(without alignments)  
854.447 Million cell updates/sec

Title: US-10-671-242-22  
Perfect score: 58  
Sequence: 1 EDDDPDGLG 10  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	75.9	406	1 A31050	calsequestrin prec
2	44	75.9	945	2 JN0901	endopeptidase Clp
3	41	70.7	420	1 S22418	calsequestrin prec
4	41	70.7	510	2 E84347	glycerol kinase [i
5	39	67.2	539	2 D71260	hypothetical prote
6	38	65.5	232	2 G75608	hypothetical prote
7	38	65.5	472	2 T45219	regulator of pre-m
8	38	65.5	526	2 T51476	hypothetical prote
9	38	65.5	784	2 D84386	ATP-dependent RNA
10	38	65.5	1047	2 T49425	hypothetical prote
11	37	63.8	127	2 AH3491	hypothetical cytos
12	37	63.8	186	2 AD3306	periplasmic dipept
13	37	63.8	228	2 AG3480	25K outer-membrane
14	37	63.8	243	2 H90015	hypothetical prote
15	37	63.8	380	1 S74913	sensory transducti
16	37	63.8	428	2 AH0709	succinylglutamic s
17	37	63.8	450	2 E82973	conserved hypothet
18	37	63.8	516	1 DACHA	procollagen-prolin
19	37	63.8	533	2 B56110	tyrosine phosphopr
20	37	63.8	584	2 S51882	topoisomerase I-re
21	37	63.8	953	2 T21292	hypothetical prote
22	37	63.8	1121	2 T52631	1-phosphatidylinos
23	36	62.1	111	2 T04405	phytochrome C - ba
24	36	62.1	151	2 T33918	hypothetical prote
25	36	62.1	153	2 A41510	myosin regulatory
26	36	62.1	215	2 AC1401	adenylate kinases
27	36	62.1	236	2 AI3377	25K outer-membrane
28	36	62.1	265	2 T46173	hypothetical prote
29	36	62.1	271	2 T42311	hypothetical prote

ALIGNMENTS

RESULT 1

A31050  
calsequestrin precursor, cardiac and skeletal muscle - chicken  
N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-1  
C;Species: Gallus gallus (Chicken)  
C;Date: 16-Jul-1999 #sequence revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: A31050; B31049; A35709; A34652  
R;Clegg, D.O.; Helder, J.C.; Hann, B.C.; Hall, D.E.; Reichardt, L.F.  
J. Cell Biol. 107, 699-705, 1988  
A;Title: Amino acid sequence and distribution of mRNA encoding a major skeletal muscle la  
domain.  
A;Reference number: A31050; MUID:883331074; PMID:3417769  
A;Accession: A31050  
A;Molecule type: mRNA  
A;Residues: 1-406 <CLE>  
A;Cross-references: UNIPROT:P19204; EMBL:Y00789; NID:g63561; PIDN:CAA68743.1; PID:g63562  
R;Hall, D.E.; Frazer, K.A.; Hann, B.C.; Reichardt, L.F.  
J. Cell Biol. 107, 687-697, 1988  
A;Title: Isolation and characterization of a laminin-binding protein from rat and chick n  
A;Reference number: A92751; MUID:883331073; PMID:3417768  
A;Accession: B31049  
A;Molecule type: protein  
A;Residues: 20-39 <HAL>  
R;Yazaki, P.J.; Salvatori, S.; Dahms, A.S.  
Biochem. Biophys. Res. Commun. 170, 1089-1095, 1990  
A;Title: Amino acid sequence of chicken calsequestrin deduced from c DNA: comparison of c  
A;Reference number: A35709; MUID:90358803; PMID:2390076  
A;Accession: A35709  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 20-147, 'I', 149-406 <YAZ>  
A;Cross-references: GB:M58048; NID:9211496; PIDN:AAA48674.1; PID:g211497  
R;Yazaki, P.J.; Salvatori, S.; Sabbadini, R.A.; Dahms, A.S.  
Biochem. Biophys. Res. Commun. 166, 898-903, 1990  
A;Title: Calsequestrin, and intracellular calcium-binding protein of skeletal muscle sarc  
ix.  
A;Reference number: A34652; MUID:90147805; PMID:2302244  
A;Accession: A34652  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 20-39 <YAZ>  
C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protein  
C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to c  
C;Superfamily: calsequestrin  
C;Keywords: calcium binding; cardiac muscle; glycoprotein; heart; skeletal muscle  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-406/Product: calsequestrin, cardiac and skeletal muscle #status experimental <NAT>  
F;335/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 75.9%; Score 44; DB 1; Length 406;  
Best Local Similarity 87.5%; Pred. No. 8.3;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

probable transposase  
calsequestrin prec  
hypothetical prote  
protoporphyryn IX  
imidazolonepropion  
hypothetical prote  
pancreatic lipase-  
lipase, CoPL-RP2 -  
probable protein k  
triacylglycerol li  
triacylglycerol li  
hypothetical prote  
ubiquitin-like fus  
transcription init  
calpain (EC 3.4.22

QY 1 EDDDDPGF 8  
 Db 275 EDDDDPGF 282

RESULT 2  
 JN0901  
 endopeptidase Clp ATP-binding chain C - Arabidopsis thaliana  
 N;Alternate names: ATP-dependent Clp proteinase regulatory chain; ERD1 protein  
 N;Contains: adenosinetriphosphatase (EC 3.6.1.3)  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 09-Jul-2004  
 C;Accession: JN0901  
 R;Kiyosue, T.; Yamaguchi-Shinozaki, K.; Shinozaki, K.  
 Biochem. Biophys. Res. Commun. 196, 1214-1220, 1993  
 A;Title: Characterization of cDNA for a dehydration-inducible gene that encodes a Clp A,  
 A;Reference number: JN0901, MUID:94071876; PMID:7504470  
 A;Accession: JN0901  
 A;Molecule type: mRNA  
 A;Residues: 1-945 <XIY>  
 A;Cross-references: UNIPROT:P42762; GB:D17582; NID:G443696; PIDN:EAA04506.1; PID:G497629  
 A;Note: This protein is homologous to the ATP-binding subunit of ATP-dependent Clp prote  
 C;Comment: This protein contains a putative chloroplast targeting signal at the amino-te  
 C;Comment: This protein interacts with a Clp-like protein and functions in the degradati  
 C;Genetics:  
 A;Gene: ERD1  
 C;Function:  
 A;Description: allows clpp to hydrolyze polypeptides and proteins, probably by a chaper  
 e activity; ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller  
 C;Superfamily: endopeptidase Clp ATP-binding chain  
 C;Keywords: ATP; GTP binding; hydrolase; molecular chaperone; nucleotide binding; P-loop  
 F;316-323/Region: nucleotide-binding motif A (P-loop)  
 F;383-388/Region: nucleotide-binding motif B  
 F;564-671/Region: nucleotide-binding motif A (P-loop)  
 F;732-737/Region: nucleotide-binding motif B  
 F;322/Binding site: ATP (Lys) #status predicted  
 F;670/Binding site: ATP (Lys) #status predicted

Query Match 75.9%; Score 44; DB 2; Length 945;  
 Best Local Similarity 80.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10  
 Db 118 EDRDQPGFLG 127

RESULT 3  
 S22418  
 calsequestrin precursor, skeletal muscle - edible frog  
 N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-  
 C;Species: Rana esculenta (edible frog)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C;Accession: S22418  
 R;Reves, S.; Vilsen, B.; Chiozzi, P.; Andersen, J.P.; Zorzato, F.  
 Biochem. J. 283, 767-772, 1992  
 A;Title: Molecular cloning, functional expression and tissue distribution of the cDNA en  
 A;Reference number: S22418, MUID:92272676; PMID:1375450  
 A;Accession: S22418  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-420 <TRE>  
 A;Cross-references: UNIPROT:P31231; EMBL:X64324; NID:G64276; PIDN:CAA45609.1; PID:G64277  
 C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protei  
 C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to c  
 C;Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in vi  
 C;Superfamily: calsequestrin  
 C;Keywords: calcium binding; glycoprotein; skeletal muscle  
 F;1-16/Domain: signal sequence #status predicted <SIG>  
 F;17-28/Domain: propeptide #status predicted <PRO>  
 F;23-420/Product: calsequestrin, fast skeletal muscle #status predicted <MAT>  
 F;338/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.7%; Score 41; DB 1; Length 420;  
 Best Local Similarity 75.0%; Pred. No. 27;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDDPGF 8  
 Db 278 EDDDDPGY 285

RESULT 4  
 E84347  
 glycerol kinase [imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C;Accession: E84347  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A;Title: Genome sequence of Halobacterium species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: E84347  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-510 <STO>  
 A;Cross-references: UNIPROT:Q9HNS5; GB:AE004437; NID:gl0581402; PIDN:AAG20145.1; GSPDB:G3  
 C;Genetics:  
 A;Gene: GlpK  
 C;Superfamily: xylulokinase

Query Match 70.7%; Score 41; DB 2; Length 510;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDDGFLG 10  
 Db 238 DDDGFLG 244

RESULT 5  
 D71260  
 hypothetical protein TP0969 - syphilis spirochete  
 C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C;Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
 C;Accession: D71260  
 R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A;Reference number: A71250; MUID:98332770; PMID:9665876  
 A;Accession: D71260  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-539 <COL>  
 A;Cross-references: UNIPROT:O83935; GB:AE001264; GB:AE000520; NID:g3323278; PIDN:AAC6592;  
 A;Experimental source: strain Nichols  
 C;Genetics:  
 A;Gene: TP0969

Query Match 67.2%; Score 39; DB 2; Length 539;  
 Best Local Similarity 77.8%; Pred. No. 77;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDDDDPGFL 9  
 Db 274 EDDRDGFL 282

RESULT 6  
 G75608

```

hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T51476
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <WHI>
A:Cross-references: UNIPROT:Q9RZ21; GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF1231
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0132
A:Map position: 2

Query Match 65.5%; Score 38; DB 2; Length 232;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DDPDGFL 9
DB 201 DDPDGFLV 207

RESULT 7
T5219
regulator of pre-meiotic DNA synthesis repl - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45219; T40109
R: Sugiyama, A.; Tanaka, K.; Okazaki, K.; Nojima, H.; Okayama, H.
EMBO J. 13, 1881-1887, 1994
A:Title: A zinc finger protein controls the onset of premeiotic DNA synthesis of fission
A:Reference number: Z22942; MUID:94222035; PMID:8168486
A:Accession: T45219
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-472 <SUG>
A:Cross-references: UNIPROT:P40379; EMBL:X77518; NID:g505307; PIDN:CRA45653.1; PID:g5053
A:Experimental source: strain L972h(-)
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21906
A:Accession: T40109
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-472 <WOO>
A:Cross-references: EMBL:AL031788; PIDN:CAA21164.1; GSPDB:GN00067; SPDB:SPBC2D10.06
A:Experimental source: strain 972h-; cosmid c2D10
C:Genetics:
A:Gene: repl; SPBC2D10.06
A:Map position: 2
C:Function:
A:Description: required for the initiation of premeiotic DNA synthesis
C:Superfamily: Schizosaccharomyces pombe regulator of pre-meiotic DNA synthesis repl

Query Match 65.5%; Score 38; DB 2; Length 472;
Best Local Similarity 87.5%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 DDPDGFLG 10
DB 458 DDGDFGLG 465

RESULT 8
T51476
hypothetical protein K3M16_70 - Arabidopsis thaliana

```

```

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: T51476
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51476
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <SAT>
A:Cross-references: UNIPROT:Q9LP52; EMBL:AL391150
A:Experimental source: cultivar Columbia; BAC clone K3M16
C:Genetics:
A:Map position: 5
A:Introns: 135/3; 218/2; 327/2; 393/1
A:Note: K3M16_70

Query Match 65.5%; Score 38; DB 2; Length 526;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDPDGFLG 10
DB 163 NDDPDAPFG 171

RESULT 9
D84386
ATP-dependent RNA helicase homolog eIF-4A [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84386
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leitchauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: D84386
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-784 <STO>
A:Cross-references: UNIPROT:Q9HWM5; GB:AE004437; NID:g10581766; PIDN:AAG20456.1; GSPDB:G
C:Genetics:
A:Gene: eif4a
C:Superfamily: ATP-dependent RNA helicase eIF-4A

Query Match 65.5%; Score 38; DB 2; Length 784;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDPDGFLG 10
DB 274 DDDSDGYQG 283

RESULT 10
T49425
hypothetical protein B17C10.60 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49425
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49425
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1047 <SCH>
A:Cross-references: UNIPROT:Q9P6B3; EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.60
A:Experimental source: BAC clone B17C10; strain OR74A
C:Genetics:

```

A:Gene: NCSP:B17C10.60  
A:Map position: 6  
A:Introns: 42/1; 55/1; 94/2; 126/3; 336/1; 489/2  
C:Superfamily: Saccharomyces cerevisiae hypothetical protein YMR196w

Query Match 65.5%; Score 38; DB 2; Length 1047;  
Best Local Similarity 77.8%; Pred. No. 2.3e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDPDGL 9  
Db 183 EDDDPDELL 191

RESULT 11  
AH3491  
hypothetical cytosolic protein BMEI1918 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AH3491  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, R.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AH3491  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-127 <KUR>  
A:Cross-references: UNIPROT:Q8YEG0; UNIPROT:Q8G3C5; GB:AE008917; PIDN:AA153099.1; PID:gl16M  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI1918  
A:Map position: I

Query Match 63.8%; Score 37; DB 2; Length 127;  
Best Local Similarity 70.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EDDDPDGL 10  
Db 108 EDDVSGILG 117

RESULT 12  
AD3306  
periplasmic dipeptide transport protein precursor BMEI0434 [imported] - Brucella melitensis  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AD3306  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, R.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AD3306  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-186 <KUR>  
A:Cross-references: UNIPROT:Q8YI10; GB:AE008917; PIDN:AA151615.1; PID:gl17982341; GSPDB:G  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0434  
A:Map position: I

Query Match 63.8%; Score 37; DB 2; Length 186;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDPDGL 10  
Db 87 DNGDPDNLG 96

RESULT 13  
AG3480  
25K outer-membrane immunogenic protein precursor [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AG3480  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, R.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AG3480  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-228 <KUR>  
A:Cross-references: UNIPROT:Q8YEP9; UNIPROT:Q8G333; GB:AE008917; PIDN:AA153010.1; PID:gl16M  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI1829  
A:Map position: I

Query Match 63.8%; Score 37; DB 2; Length 228;  
Best Local Similarity 77.8%; Pred. No. 67;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDDPDGL 10  
Db 69 DIKPDGL 77

RESULT 14  
H90015  
hypothetical protein SA1999 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: H90015  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; PMID:11418146  
A:Accession: H90015  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-243 <KUR>  
A:Cross-references: UNIPROT:Q99S72; GB:BA000018; PID:gl13701997; PIDN:BA043289.1; GSPDB:G  
A:Experimental source: strain N315  
C:Genetics:  
A:Superfamily: conserved hypothetical protein b1120  
A:Gene: SA1999

Query Match 63.8%; Score 37; DB 2; Length 243;  
Best Local Similarity 62.5%; Pred. No. 72;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DDDPDGL 9  
Db 62 EDDPDGL 69

RESULT 15  
S74913  
sensory transduction system regulatory protein slr1400 - Synechocystis sp. (strain PCC 6803)  
N:Alternate names: protein slr1400  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S74913  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996



A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.  
A;Reference number: S74322; MUID:97061201; PMID:8905231  
A;Accession: S74913  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-380 <KAN>  
A;Cross-references: UNIPROT:P72936; EMBL:D90902; GS:AB001339; NID:g1652027; PIDN:BAA1695  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C;Superfamily: sensory transduction system regulatory protein; response regulator homolog  
C;Keywords: phosphoprotein  
F;6-117/Domain: response regulator homology <RRH>  
F;54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 63.8%; Score 37; DB 1; Length 380;  
Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DDDPDGF 8  
| | | | |  
Db 10 DDDPDNF 16

Search completed: September 20, 2005, 12:43:23  
Job time : 2.12607 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:00:18 ; Search time 5.07956 Seconds  
(without alignments)  
1008.117 Million cell updates/sec

Title: US-10-671-242-22  
Perfect score: 58  
Sequence: 1 EDDDPDGLG 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	58	100.0	236	2	Q96IM4	Q6p0r3 homo sapien
2	58	100.0	285	2	Q6P0R3	Q6p0r3 homo sapien
3	58	100.0	295	2	Q6PIS7	Q6p1s7 homo sapien
4	58	100.0	538	2	Q6SPF0	Q6spfo homo sapien
5	54	93.1	550	2	Q6SPE9	Q6spe9 oryctolagus
6	46	79.3	872	2	Q6CYD6	Q6cyd6 kluyveromyc
7	45	77.6	509	2	Q75DH4	P19204 gallus gall
8	44	75.9	406	1	CAQ1CHICK	Q75dh4 ashbya gos
9	44	75.9	640	2	Q94C10	Q94c10 arabidopsis
10	44	75.9	945	1	ERD1ARATH	P42762 arabidopsis
11	43	74.1	340	2	Q81XQ4	Q81xq4 homo sapien
12	43	74.1	346	2	Q8K2V8	Q8k2v8 mus musculu
13	43	74.1	351	2	Q9C0F9	Q9c0f9 homo sapien
14	43	74.1	364	2	Q69ZC8	Q69zc8 mus musculu
15	43	74.1	804	2	Q9VWD3	Q9vwd3 drosophila
16	41	70.7	273	2	Q6A9M3	Q6a9m3 propionibac
17	41	70.7	293	2	Q64RB7	Q64rb7 bacteroides
18	41	70.7	300	2	Q6PBF1	Q6pbf1 xenopus tro
19	41	70.7	354	2	Q96EM0	Q96em0 homo sapien
20	41	70.7	354	2	Q96LJ5	Q96lj5 homo sapien
21	41	70.7	396	2	Q6TU32	Q6tu32 aplysia cal
22	41	70.7	408	2	Q6AR41	Q6ar41 desulfotale
23	41	70.7	420	1	CAQ1_RANES	F31231 rana esculu
24	41	70.7	428	2	Q7ZWM0	Q7zwm0 xenopus lae
25	41	70.7	435	2	Q6ZUV1	Q6zuv1 homo sapien
26	41	70.7	450	2	Q7F9F0	Q7f9f0 oryza sativ
27	41	70.7	510	1	GLPK HALN1	Q9hns5 halobacteri
28	40	69.0	265	2	Q6WL19	Q6wl19 rhizobium s
29	40	69.0	316	2	Q82GF4	Q82gf4 streptomyce
30	40	69.0	324	2	Q9KXV9	Q9kxv9 streptomyce
31	40	69.0	409	2	Q6D116	Q6d116 brachydanio

32 40 69.0 623 2 Q7PUC7  
33 40 69.0 673 2 Q653C5  
34 40 69.0 729 2 Q9GSZ5  
35 40 69.0 859 2 Q6CL30  
36 40 69.0 1167 2 Q8P2I8  
37 40 69.0 1311 2 Q9VXJ5  
38 39 67.2 77 2 Q8W1M3  
39 39 67.2 84 2 Q8LP75  
40 39 67.2 89 2 Q8LP76  
41 39 67.2 90 2 Q8LP77  
42 39 67.2 96 2 Q8LP79  
43 39 67.2 99 2 Q8W1M4  
44 39 67.2 108 2 Q83148  
45 39 67.2 254 2 Q7NCD0

## ALIGNMENTS

RESULT 1  
Q96IM4 PRELIMINARY; PRT; 236 AA.  
AC Q96IM4;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE LOC90378 protein (Hypothetical protein) (Fragment).  
GN Name=LOC90378;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Director MGC Project;  
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC007384; AAH07384.2; -;  
DR EMBL; BC080588; AAH080588.1; -;  
DR HSSP; P39769; 1KW4  
DR InterPro; IPR001660; SAM.  
DR InterPro; IPR001993; SAM\_homology.  
DR Pfam; PF00536; SAM\_1; 1.  
DR SMART; SM00454; SAM; 1.

Q7puc7 anopheles g  
Q653c5 oryza sativ  
Q9gsz5 drosophila  
Q6cl30 kluyveromyc  
Q8p2i8 dictyosteli  
Q9vxj5 drosophila  
Q8w1m3 solanum cha  
Q8lp75 lycopersico  
Q8lp76 lycopersico  
Q8lp77 lycopersico  
Q8lp79 lycopersico  
Q8w1m4 solanum cha  
Q83148 murid herpe  
Q7ncd0 gloeobacter

```

DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 236 AA; 25338 MW; AECF03B5D165FC0E CRC64;

Query Match 100.0%; Score 58; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 227 EDDDPDGFGLG 236

RESULT 2
Q6P0R3 PRELIMINARY; PRT; 285 AA.
ID Q6P0R3
AC Q6P0R3
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065477; AH65477.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 285 AA; 30465 MW; FF2F936CAF1F901 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 276 EDDDPDGFGLG 285

RESULT 3
Q6PIS7 PRELIMINARY; PRT; 295 AA.
ID Q6PIS7
AC Q6PIS7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
DE LOC90378 protein (Fragment).
GN Name=LOC90378;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030129; AAH30129.1; -.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR010993; SAM_homology.
DR Pfam; PF00536; SAM 1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 31502 MW; F96F943C7A696F19 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10
Db 286 EDDDPDGFGLG 295

RESULT 4
Q6SPF0 PRELIMINARY; PRT; 538 AA.
ID Q6SPF0
AC Q6SPF0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453840; AAR24087.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 538 AA; 56051 MW; 083A170790654F2F CRC64;

Query Match 100.0%; Score 58; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db |||||
529 EDDDPDGFLG 538

RESULT 5
Q6SPE9 PRELIMINARY; PRT; 550 AA.
AC Q6SPE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Atherin.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Lees A.M., Deconinck A.E., Campbell B.D., Lees R.S.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY453841; AAR24088.1; -.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00536; SAM_1; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 550 AA; 57019 MW; AF1CF2B780DB79A7 CRC64;

Query Match 93.18; Score 54; DB 2; Length 550;
Best Local Similarity 90.0%; Pred. No. 2.1;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db |||||
541 EDDDPDGFLG 550

RESULT 6
Q6CYD6 PRELIMINARY; PRT; 872 AA.
AC Q6CYD6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P53076 Saccharomyces cerevisiae YGL227w VID30
DE singleton.
GN ORFNames=KLLA0A01221g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382121; CAH02641.1; -.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR InterPro; IPR006595; CTLL_C.
DR InterPro; IPR006594; LISH.
DR InterPro; IPR000342; RGS.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00622; SPRY; 1.
DR SMART; SM00449; SPRY; 1.
DR PROSITE; PS50897; CTLL; 1.
DR PROSITE; PS50896; LISH; 1.
SQ SEQUENCE 872 AA; 99640 MW; 8FACE92753B2DC8 CRC64;

Query Match 79.3%; Score 46; DB 2; Length 872;
Best Local Similarity 88.9%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EDDDPDGFL 9
Db |||||
589 EDDDPDGFL 597

RESULT 7
Q75DH4 PRELIMINARY; PRT; 509 AA.
AC Q75DH4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ABR052Wp.
GN ORFNames=ABR052W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegeli S.E., Dietrich F.S., Brachat S., Lerch A., Gaffney T.,
Philippe P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AR016815; AAS50822.1; -.
DR AGD; ABR052W; -.
DR GO; GO:0016742; F:hydroxymethyl-, formyl- and related transfe. .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR002376; Formyl transf N.
SQ SEQUENCE 509 AA; 56603 MW; A66BAC28DD62FF88 CRC64;

Query Match 77.6%; Score 45; DB 2; Length 509;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFL 8
Db |||||

```

```

Db          302 EDDDPDGY 309

RESULT 8
CAQ1_CHICK STANDARD; PRT; 406 AA.
AC PI9204;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 03-JUL-2004 (Rel. 44, Last annotation update)
DE Calsequestrin, skeletal muscle isoform precursor (Calsequestrin 1)
DE (Aspartactin) (Laminin-binding protein).
GN Name=CASQ1;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCB1_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=88331074; PubMed=3417769; DOI=10.1083/jcb.107.2.699;
RA Clegg D.O., Heider J.C., Hann B.C., Hall D.E., Reichardt L.F.;
RT "Amino acid sequence and distribution of mRNA encoding a major
RT skeletal muscle laminin binding protein: an extracellular matrix-
RT associated protein with an unusual COOH-terminal polyaspartate
RT domain.";
RL J. Cell Biol. 107:699-705(1998).
RN [2]
RP SEQUENCE OF 20-406 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90358803; PubMed=2390076;
RA Yazaki P.J., Salvatori S., Dahms A.S.;
RT "Amino acid sequence of chicken calsequestrin deduced from cDNA:
RT comparison of calsequestrin and aspartactin.";
RL Biochem. Biophys. Res. Commun. 170:1089-1095(1990).
RN [3]
RP ERATUM.
RX MEDLINE=91083667; PubMed=2260982;
RA Yazaki P.J., Salvatori S., Dahms A.S.;
RL Biochem. Biophys. Res. Commun. 173:763-763(1990).
RN [4]
RP SEQUENCE OF 20-39.
RC STRAIN=White leghorn; TISSUE=Skeletal muscle;
RX MEDLINE=88331073; PubMed=3417768; DOI=10.1083/jcb.107.2.687;
RA Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;
RT "Isolation and characterization of a laminin-binding protein from rat
RT and chick muscle.";
RL J. Cell Biol. 107:687-697(1998).
RN [5]
RP SEQUENCE OF 20-39.
RC STRAIN=New Hampshire;
RX MEDLINE=90147805; PubMed=2302244;
RA Yazaki P.J., Salvatori S., Sabbadini R.A., Dahms A.S.;
RT "Calsequestrin, an intracellular calcium-binding protein of skeletal
RT muscle sarcoplasmic reticulum, is homologous to aspartactin, a
RT putative laminin-binding protein of the extracellular matrix.";
RL Biochem. Biophys. Res. Commun. 166:898-903(1990).
CC -1- FUNCTION: Calsequestrin is a high-capacity, moderate affinity,
CC calcium-binding protein and thus acts as an internal calcium store
CC in muscle. The release of calcium bound to calsequestrin through a
CC calcium release channel triggers muscle contraction. Binds 40 to
CC 50 moles of calcium. Also binds laminin.
CC -1- SUBCELLULAR LOCATION: This isoform of calsequestrin occurs in the
CC sarcoplasmic reticulum's terminal cisternae luminal spaces of fast
CC skeletal muscle cells. Aspartactin is found in the basal lamina
CC surrounding individual muscle fibers.
CC -1- TISSUE SPECIFICITY: Skeletal and heart muscle.
CC -1- SIMILARITY: Belongs to the calsequestrin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

---

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Y00789; CAA68743.1; -
DR EMBL; M58048; AAA48674.1; -
DR PIR; A31050; A31050.
DR HSSP; P07221; 1A8V.
DR InterPro; IPR001393; Calsequestrin.
DR Pfam; PF01216; Calsequestrin; 1.
DR PRINTS; PR00312; CALSEQUESTRN.
DR PROSITE; PS00863; CALSEQUESTRN_1; 1.
DR PROSITE; PS00864; CALSEQUESTRN_2; 1.
KW Calcium-binding; Direct protein sequencing; Glycoprotein;
KW Muscle protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 406 Calsequestrin, skeletal muscle isoform.
FT DOMAIN 374 406 Poly-Asp.
FT CARBOHYD 335 335 N-linked (GlcNAc...) (Probable).
FT CONFLICT 148 148 V -> I (in Ref. 4).
SQ SEQUENCE 406 AA; 47151 MW; 928CFS4785A2DC7C CRC64;

Query Match 75.9%; Score 44; DB 1; Length 406;
Best Local Similarity 87.5%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGF 8
DB 275 EDDDPDGF 282

RESULT 9
Q94C10 PRELIMINARY; PRT; 640 AA.
ID Q94C10;
AC Q94C10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AT5G51070/K3K7_27.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
RA Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY037264; AAK59865.1; -
DR HSSP; Q9RA63; IQVR.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0019538; P:protein metabolism; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR001270; Chaprinin_clpA/B.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02861; Clp_N; 2.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00870; CLPAB_1; UNKNOWN_1.
KW ATP-BINDING.
SQ SEQUENCE 640 AA; 69610 MW; 4C477ECA608E84BB CRC64;

Query Match 75.9%; Score 44; DB 2; Length 640;

```

Best Local Similarity 80.0%; Pred. No. 1e+02; Mismatches 0; Indels 2; Gaps 0;  
Matches 8; Conservative 0;

QY 1 EDDDDPGFLG 10  
DB 118 EDRDPQFLG 127

## RESULT 10

ERD1\_ARATH STANDARD; PRT; 945 AA.  
AC P42762;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ERD1 protein, chloroplast precursor.  
GN Name=ERD1; OrderedLocNames=At5g1070; ORFNames=K3K7.27;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=94071876; PubMed=7504470;  
RA Kiyose T., Yamaguchi-Shinozaki K., Shinozaki K.;  
RT "Characterization of cDNA for a dehydration-inducible gene that  
RT encodes a Clp A, B-like protein in Arabidopsis thaliana L.";  
RL Biochem. Biophys. Res. Commun. 196:1214-1220(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20181125; PubMed=10718197;  
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
RT clones";  
RL DNA Res. 7:31-63(2000).  
CC -I- FUNCTION: May interact with a clpp-like protease involved in  
CC degradation of denatured proteins in the chloroplast.  
CC -I- SUBCELLULAR LOCATION: Chloroplast (Potential).  
CC -I- INDUCTION: By dehydration stress. Induced after one hour of  
CC dehydration-stress and reaches maximal levels after 10 hours.  
CC -I- SIMILARITY: Belongs to the clpA/clpB family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC ENBL; D17582; BAA04506.1; --  
CC ENBL; AB023044; BAB10330.1; --  
CC ENBL; AB017063; BAB10330.1; JOINED.  
CC FR; JN0901; JN0901.  
CC HSP; P03815; IJBK.  
CC GeneFarm; 2731; 97.  
CC InterPro; IPR003593; AAA\_ATPase.  
CC InterPro; IPR003959; AAA\_ATPase\_cent.  
CC InterPro; IPR001270; Chaprinin\_clpA/B.  
CC InterPro; IPR004176; Clp\_N.  
CC Pfam; PF00004; AAA; 1.  
CC PRINTS; PR00361; Clp\_N; 2.  
CC SMART; SM00382; AAA; 2.  
CC PROSITE; PS00870; CLPAB\_1; 1.  
CC PROSITE; PS00871; CLPAB\_2; 1.  
KW ATP-binding; Chaperone; Chloroplast; Repeat; Transit peptide.  
FT TRANSIT 1 ? Chloroplast (potential).

FT CHAIN ? 945 ERD1 protein.  
FT DOMAIN 271 523 I.  
FT DOMAIN 590 781 II.  
FT NP\_BIND 316 323 ATP (Potential).  
FT NP\_BIND 664 671 ATP (Potential).  
SQ SEQUENCE 945 AA; 103234 MW; 81EF2332C78F656B CRC64;

Query Match 75.9%; Score 44; DB 1; Length 945;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10  
DB 118 EDRDPQFLG 127

## RESULT 11

Q81XQ4 PRELIMINARY; PRT; 340 AA.  
AC Q81XQ4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE KIAA1704.  
GN Name=KIAA1704;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Strausberg R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR ENBL; BC039586; AAH39586.1; --  
SQ SEQUENCE 340 AA; 38141 MW; F752A6A2917132E6 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 340;  
Best Local Similarity 70.0%; Pred. No. 78;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDPGFLG 10  
DB 83 DDDDDGDFG 92

## RESULT 12

Q8K2V8 PRELIMINARY; PRT; 346 AA.  
ID Q8K2V8

AC Q8K2V8;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DT 23-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE RIKEN cDNA 120001118 (Mus musculus 7 days embryo whole body cDNA,  
 DE RIKEN full-length enriched library, clone:C430016019 product:similar  
 DE to PROTEIN PHOSPHATASE 4 REGULATORY SUBUNIT 2).  
 GN Name:120001118Rik;  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RX Strausberg R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX STRAIN=C57BL/6J; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [7]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito D., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki R., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC029731; AAH29731.1; -;  
 DR EMBL; AK082912; BAC38684.1; -;  
 DR MGD; MGI:1914717; 120001118Rik.  
 SQ SEQUENCE 346 AA; 38954 MW; F7EBC0CCD8267633 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 346;  
 Best Local Similarity 70.0%; Pred. No. 80;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDGDFLG 10  
 Db 82 DDDDDGDFFG 91

RESULT 13

Q9COF9 PRELIMINARY; PRT; 351 AA.

AC Q9COF9;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE KIAA1704 protein (Fragment).  
 GN Name:KIAA1704;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082932; PubMed=11214970;  
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:347-355(2000).  
 DR EMBL; AB051491; BAB21795.1; -;  
 FT NON TER 1  
 SQ SEQUENCE 351 AA; 39406 MW; 6F5252CB5A6FE2F2 CRC64;

Query Match 74.1%; Score 43; DB 2; Length 351;  
 Best Local Similarity 70.0%; Pred. No. 81;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EDDDDGDFLG 10  
 :|||||



```

Db          94 DDDDDGFFG 103
RESULT 14
Q69ZC8      PRELIMINARY;      PRT;      364 AA.
AC          Q69ZC8;
AT          25-OCT-2004 (T-EMBLrel. 28, Created)
DT          25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT          25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE          MKIAAL704 protein (Fragment).
GN          Name=MKIAAL704;
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX          NCBI_TaxID=10090;
RN          [1]
RP          SEQUENCE FROM N.A.
RA          Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA          Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA          Nagase T., Ohara O., Koga H.;
RT          "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT          IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT          cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT          Randomly Sampled from Size-Fractionated Libraries.";
RL          DNA Res. 11:205-218(2004).
DR          EMBL; AK173238; BAD32516.1; -.
FT          NON TER      1
SQ          SEQUENCE      364 AA; 40946 MW; C9C3A38CBES8D9AS CRC64;

Query Match      74.1%; Score 43; DB 2; Length 364;
Best Local Similarity 70.0%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy          1 EDDDDGFFG 10
Db          100 DDDDDGFFG 109
RESULT 15
Q9VWD3      PRELIMINARY;      PRT;      804 AA.
ID          Q9VWD3;
AC          Q9VWD3;
AT          01-MAY-2000 (T-EMBLrel. 13, Created)
DT          01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT          01-MAY-2004 (T-EMBLrel. 26, Last annotation update)
DE          CG12531-PA.
GN          ORFNames=CG12531;
OS          Drosophila melanogaster (Fruit fly).
OC          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC          Ephydroidea; Drosophiliidae; Drosophila.
OX          NCBI_TaxID=7227;
RN          [1]
RP          SEQUENCE FROM N.A.
RA          Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA          Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.P.,
RA          Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA          Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA          Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA          Abrial J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA          Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA          Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA          Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA          Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA          Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA          de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA          Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA          Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA          Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA          Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA          Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA          Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA          Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA          Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA          Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA          Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA          Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA          Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA          Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA          Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA          Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA          Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA          Spier E.C., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA          Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA          Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA          Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA          Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA          Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA          Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT          "The genome sequence of Drosophila melanogaster.";
RL          Science 287:2185-2195(2000).
RN          [2]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=22426065; PubMed=12537568;
RA          Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA          Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA          George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA          Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA          Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA          Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT          "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT          melanogaster euchromatic genome sequence.";
RL          Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN          [3]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=22426070; PubMed=12537573;
RA          Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA          Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA          Ashburner M., Celniker S.E.;
RT          "The transposable elements of the Drosophila melanogaster euchromatin:
RT          a genomics perspective.";
RL          Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN          [4]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=22426069; PubMed=12537572;
RA          Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA          Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA          Smith C.D., Tupry J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA          Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA          Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA          Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA          Lewis S.E.;
RT          "Annotation of the Drosophila melanogaster euchromatic genome: a
RT          systematic review.";
RL          Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN          [5]
RP          SEQUENCE FROM N.A.
RG          FlyBase;
RL          Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN          [6]
RP          SEQUENCE FROM N.A.
RG          FlyBase;
RL          Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR          EMBL; AE003513; AAF49012.1; -.
DR          FlyBase; FBgn0031064; CG12531.
DR          GO; GO:0016021; C:integral to membrane; IEA.
DR          GO; GO:0005279; P:amino acid-polyamine transporter activity; IEA.
DR          GO; GO:0006865; P:amino acid transport; IEA.
DR          GO; GO:0006810; P:transport; IEA.
DR          InterPro; IPR002293; AA/rel_permease.
DR          InterPro; IPR004841; Permease_region.
DR          Pfam; PF00324; AA_permease; 2.
KW          Transmembrane; Transport.

```

SQ SEQUENCE 804 AA; 86555 MW; 6D13A1DB1D2EC1AB CRC64;  
Query Match 74.1%; Score 43; DB 2; Length 804;  
Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 DDDPDGFLG 10  
|||:||||  
Db 483 DDSFEGFLG 491

Search completed: September 20, 2005, 12:41:47  
Job time : 6.07956 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 12:26:49 ; Search time 5.5202 Seconds  
(without alignments)  
733.538 Million cell updates/sec

Title: US-10-671-242-22  
Perfect score: 58  
Sequence: 1 EDDPDGFLG 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	10	9	US-09-962-055-22
2	58	100.0	10	9	US-09-976-740-22
3	58	100.0	10	13	US-10-023-529-22
4	58	100.0	10	13	US-10-023-523-22
5	58	100.0	10	15	US-10-616-187-22
6	58	100.0	10	15	US-10-671-242-22
7	58	100.0	217	9	US-09-962-055-7
8	58	100.0	217	9	US-09-976-740-7
9	58	100.0	217	13	US-10-023-529-7
10	58	100.0	217	13	US-10-023-523-7
11	58	100.0	217	15	US-10-616-187-7

58	100.0	217	15	US-10-671-242-7	Sequence 7, Appli
58	100.0	241	9	US-09-925-298-665	Sequence 665, App
58	100.0	241	14	US-10-102-806-665	Sequence 665, App
58	100.0	538	9	US-09-976-740-43	Sequence 43, Appl
58	100.0	538	13	US-10-023-529-43	Sequence 43, Appl
58	100.0	538	13	US-10-023-523-43	Sequence 43, Appl
58	100.0	538	15	US-10-616-187-43	Sequence 43, Appl
58	100.0	538	15	US-10-671-242-43	Sequence 43, Appl
54	93.1	232	9	US-09-962-055-3	Sequence 3, Appli
54	93.1	232	9	US-09-976-740-3	Sequence 3, Appli
54	93.1	232	13	US-10-023-529-3	Sequence 3, Appli
54	93.1	232	13	US-10-023-523-3	Sequence 3, Appli
54	93.1	232	15	US-10-616-187-3	Sequence 3, Appli
54	93.1	232	15	US-10-671-242-3	Sequence 3, Appli
54	93.1	252	9	US-09-962-055-4	Sequence 4, Appli
54	93.1	252	9	US-09-976-740-4	Sequence 4, Appli
54	93.1	252	13	US-10-023-529-4	Sequence 4, Appli
54	93.1	252	13	US-10-023-523-4	Sequence 4, Appli
54	93.1	252	15	US-10-616-187-4	Sequence 4, Appli
54	93.1	252	15	US-10-671-242-4	Sequence 4, Appli
54	93.1	317	9	US-09-962-055-2	Sequence 2, Appli
54	93.1	317	9	US-09-976-740-2	Sequence 2, Appli
54	93.1	317	13	US-10-023-529-2	Sequence 2, Appli
54	93.1	317	13	US-10-023-523-2	Sequence 2, Appli
54	93.1	317	15	US-10-616-187-2	Sequence 2, Appli
54	93.1	317	15	US-10-671-242-2	Sequence 2, Appli
54	93.1	550	9	US-09-976-740-47	Sequence 47, Appl
54	93.1	550	13	US-10-023-529-47	Sequence 47, Appl
54	93.1	550	13	US-10-023-523-47	Sequence 47, Appl
54	93.1	550	15	US-10-616-187-47	Sequence 47, Appl
54	93.1	550	15	US-10-671-242-47	Sequence 47, Appl
44	75.9	640	17	US-10-732-923-7094	Sequence 7094, Ap
44	75.9	945	9	US-09-812-350-3	Sequence 3, Appli
44	75.9	945	17	US-10-732-923-7090	Sequence 7090, Ap

ALIGNMENTS

RESULT 1  
US-09-962-055-22  
; Sequence 22, Application US/09962055  
; Patent No. US20020052033A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; Lees, Robert S.  
; Law, Simon W.  
; Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA: US/09/962,055  
; APPLICATION NUMBER: US/09/962,055  
; FILING DATE: 24-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/979,608  
; FILING DATE: 26-NOV-1997  
; APPLICATION NUMBER: US 60/031,930  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-962-055-22

Query Match      100.0%; Score 58; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 2
US-09-976-740-22
; Sequence 22, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-22

Query Match      100.0%; Score 58; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 3
US-10-023-529-22
; Sequence 22, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

```

```

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-22

Query Match      100.0%; Score 58; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

RESULT 4
US-10-023-523-22
; Sequence 22, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-22

Query Match      100.0%; Score 58; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
Db 1 EDDDPDGFLG 10

```

RESULT 5  
US-10-616-187-22  
; Sequence 22, Application US/10616187  
; Publication No. US20040013668A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/10/616,187  
; PRIOR FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: US/09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-616-187-22

Query Match 100.0%; Score 58; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10  
Db 1 EDDDPDGFLG 10

RESULT 6  
US-10-671-242-22  
; Sequence 22, Application US/10671242  
; Publication No. US20040040049A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/10/671,242  
; PRIOR FILING DATE: 2003-09-24  
; PRIOR APPLICATION NUMBER: US/09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 10

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-671-242-22

Query Match 100.0%; Score 58; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10  
Db 1 EDDDPDGFLG 10

RESULT 7  
US-09-962-055-7  
; Sequence 7, Application US/09962055  
; Patent No. US20020052033A1  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
; TITLE OF INVENTION: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
; TITLE OF INVENTION: TREATING ATHEROSCLEROSIS  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/962,055  
; FILING DATE: 24-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/979,608  
; FILING DATE: 26-NOV-1997  
; APPLICATION NUMBER: US 60/031,930  
; FILING DATE: 27-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 217 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-962-055-7

Query Match 100.0%; Score 58; DB 9; Length 217;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10  
Db 208 EDDDPDGFLG 217

RESULT 8  
US-09-976-740-7

```
; Sequence 7, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-740-7

Query Match      100.0%; Score 58; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDDDPDGLG 10
Db      208 EDDDPDGLG 217

RESULT 9
US-10-023-529-7
; Sequence 7, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 58; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDDDPDGLG 10
Db      208 EDDDPDGLG 217

RESULT 10
US-10-023-523-7
; Sequence 7, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-7

Query Match      100.0%; Score 58; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDDDPDGLG 10
Db      208 EDDDPDGLG 217

RESULT 11
US-10-616-187-7
; Sequence 7, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-529-7

Query Match      100.0%; Score 58; DB 13; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EDDDPDGLG 10
Db      208 EDDDPDGLG 217
```

```
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-187-7

Query Match      100.0%; Score 58; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 208 EDDDPDGFLG 217

RESULT 12
US-10-671-242-7
; Sequence 7, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-09-24
; PRIOR FILING DATE: 2000-03-02
; PRIOR FILING DATE: 2000-03-02
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1996-11-27
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-671-242-7

Query Match      100.0%; Score 58; DB 15; Length 217;
Best Local Similarity 100.0%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 208 EDDDPDGFLG 217

RESULT 13
US-09-925-298-665
; Sequence 665, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; APPLICANT: Lees, Robert S.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

Query Match      100.0%; Score 58; DB 14; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 232 EDDDPDGFLG 241

RESULT 14
US-10-102-806-665
; Sequence 665, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 665
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-665

Query Match      100.0%; Score 58; DB 9; Length 241;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFLG 10
   |||||
Db 232 EDDDPDGFLG 241

RESULT 15
US-09-976-740-43
; Sequence 43, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
```

; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; TITLE OF INVENTION: ATHEROSCLEROSIS  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/976,740  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 09/616,289  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 538  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-976-740-43

Query Match 100.0%; Score 58; DB 9; Length 538;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EDDDPDGFGLG 10  
|||||||  
Db 529 EDDDPDGFGLG 538

Search completed: September 20, 2005, 12:53:07  
Job time : 6.5202 secs